

# Lifelogs and Personal Data. An Unsolved Challenge

## Cathal Gurrin

Full Professor at the School of Computing, Dublin City University (DCU), Ireland  
Deputy Director of the national Adapt Centre for AI-driven Digital Content Technologies  
Failte Ireland Conference Ambassador of the year (2025)



Cathal Gurrin (2026)

Lifelogging is the process of passively capturing, storing, organizing, and retrieving digital traces of human life activities and experiences over time

The lifelog (dataset) can enable a wide range of positive applications for the individual from enhanced health, memory support and productive enhancement

To Motivate  
Behaviour Change



To Remember  
Past Events



To Represent &  
Support



Wellness ++

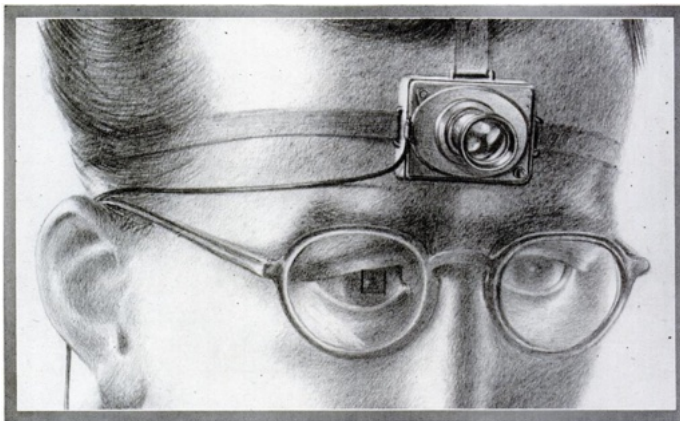


Productivity ++



Independence / Security ++

## The Memex Vision (1945)



A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SHOWS THE OBJECT

<https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>



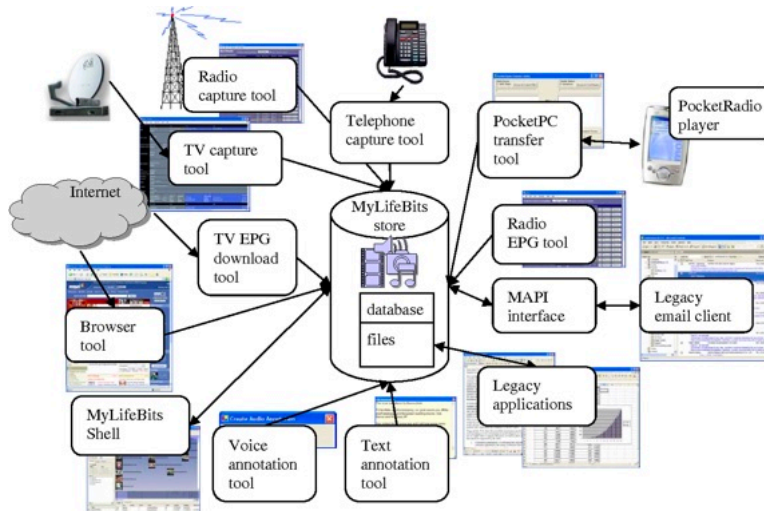
*"Consider a future device ... in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory".*

Vannevar Bush, 1945



Cathal Gurrin (2026)

## MyLifeBits (SIGIR 2004)



*Gordon Bell's effort to create a lifelog at Microsoft. The MyLifeBits project was an experiment in life-logging based on Vannevar Bush's Memex.*

Cathal Gurrin (2026)

# Modeling the Richness of Human Life

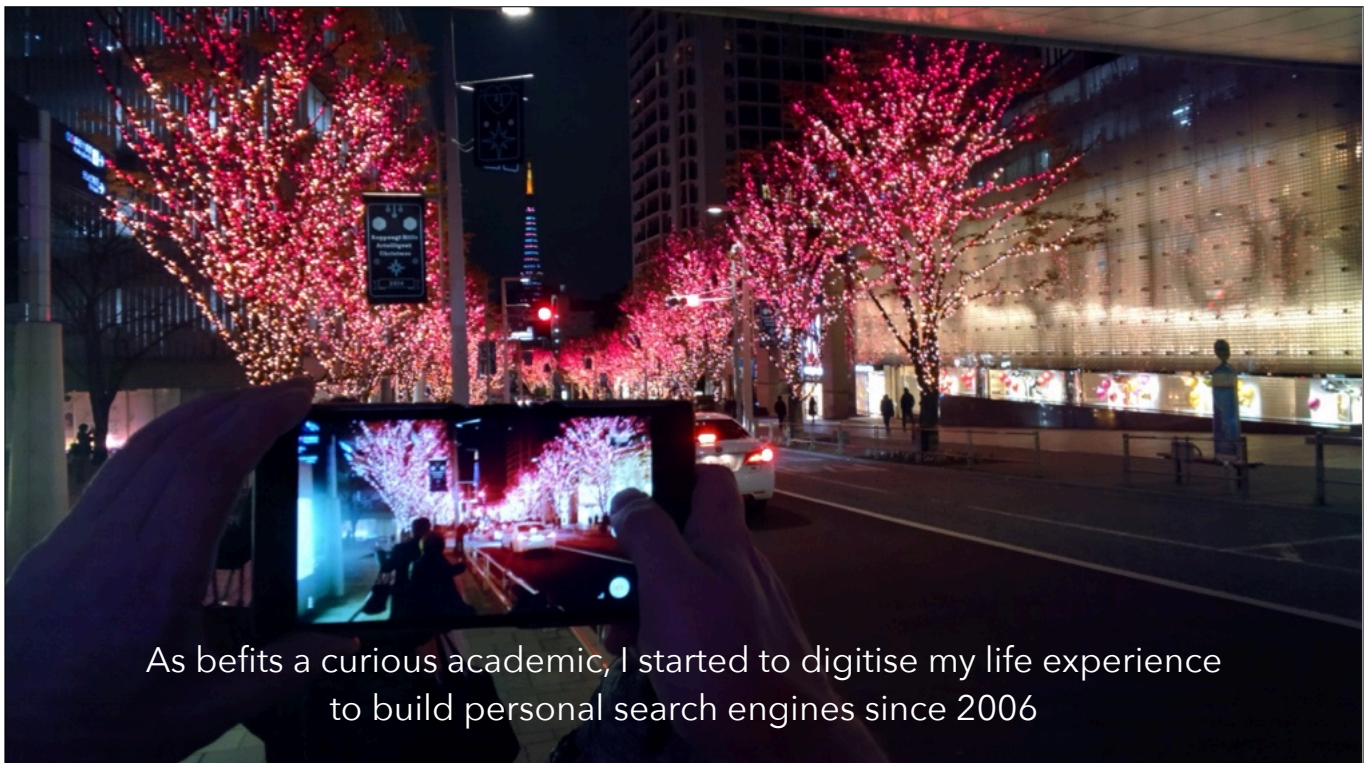
We spend our lives in complex, continuous multimodal environments. If we can capture and model these, we can provide many life-enhancing services.

Most computer vision / multimedia datasets capture short, isolated actions, whereas real daily life is:

- Long-duration
- Context-rich
- Socially interactive
- Sequential and structured

If we want AI tools to understand human life and not just isolated actions, then we need datasets **(lifelogs)** that capture life as it is actually lived. We also need to motivate people to share these datasets with the community.

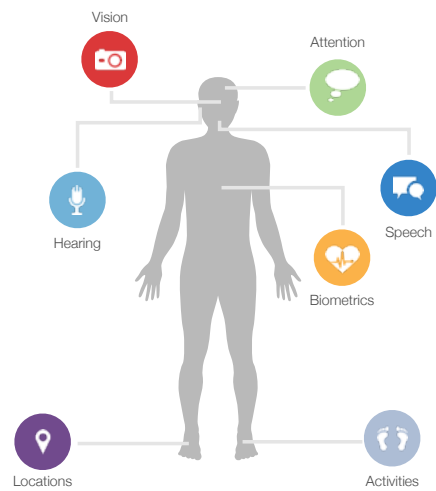
Cathal Gurrin (2026)



Driven by **low-cost, reliable, ubiquitous sensors**, massive volumes of personal lifelog data can be easily captured and stored long-term at **low storage cost**.

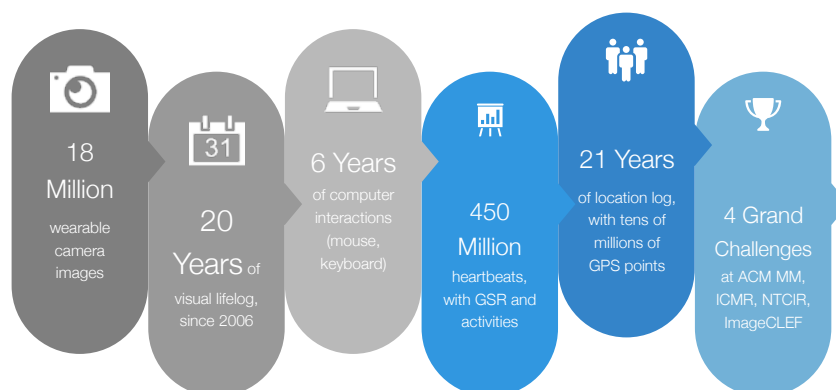
New forms of **multimedia and multimodal data analytics tools** means that we can extract value from these lifelogs.

This is what I am interested in... applying Information Retrieval tools to these lifeless



Cathal Gurrin (2026)

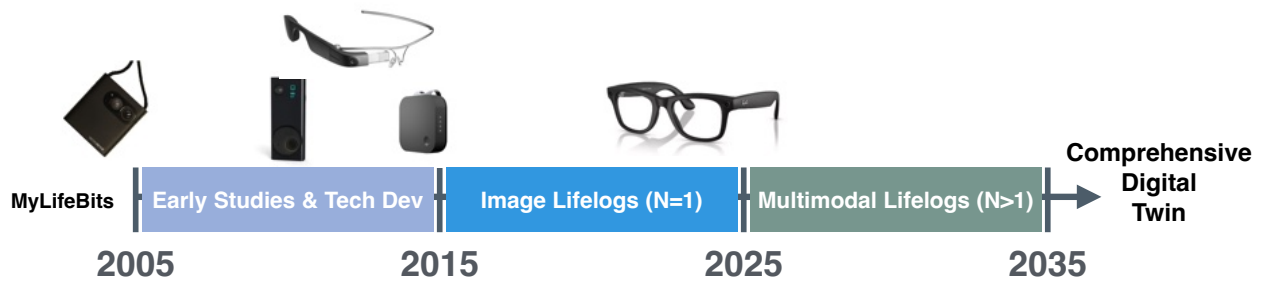
## My Lifelog in Numbers



About 1TB per year

Cathal Gurrin (2026)

I see this as a **three-decade long challenge** that brings us from the first 'lifelogging' device (the Microsoft Sensecam) through the first generation challenges (e.g. ACM Lifelog Search Challenge), into the era of Digital Twins of the entirety of life experience, captured multi modally and stored indefinitely for life enhancement.



Cathal Gurrin (2026)

## Agenda

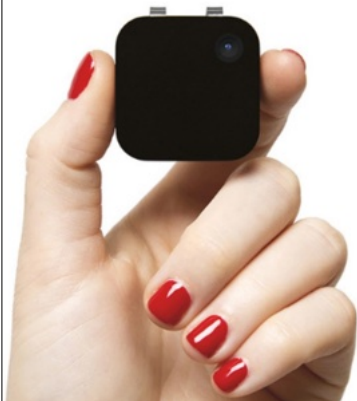
1. Early Experiments in Lifelogging
2. State-of-the-Art from Grand Challenges
3. Next Steps in Lifelog Access & Retrieval

# 1

## Early Experiments in Lifelogging

Cathal Gurrin (2026)

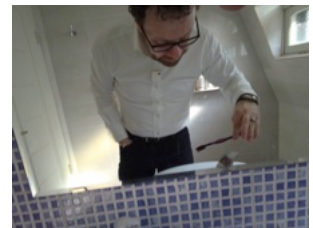
## Wearable Cameras



Capture  $N$  images per minute automatically. Some video capture, but images are often sufficient.

1,500 images per day = 1GB.

16 hours of 4K video = 1TB

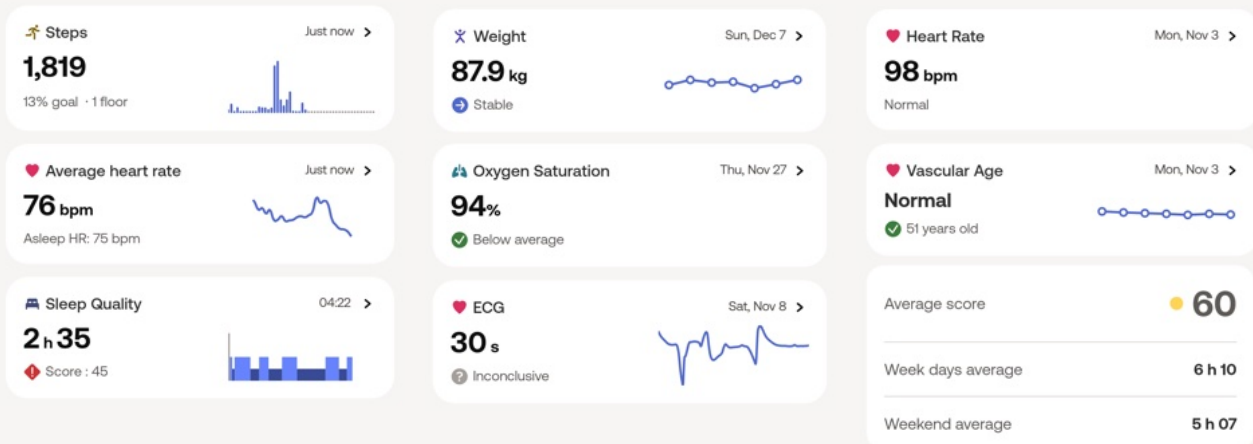


Cathal Gurrin (2026)



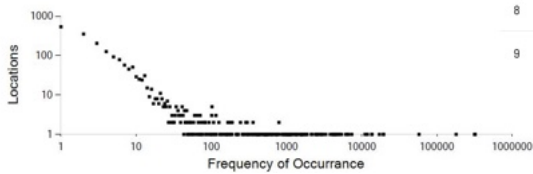
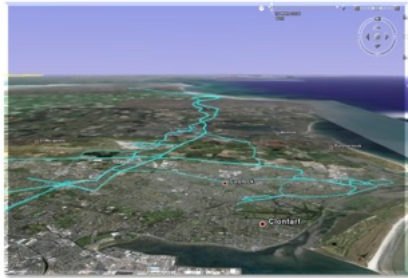
Cathal Gurrin (2026)

Biometrics are readily available and reflect our bodies natural state at any point in time.

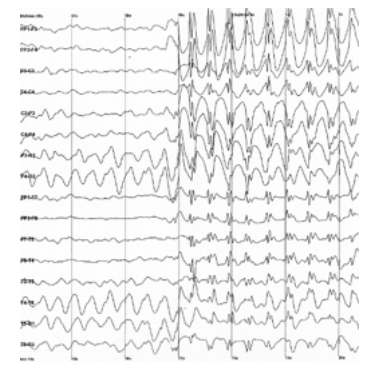


Cathal Gurrin (2026)

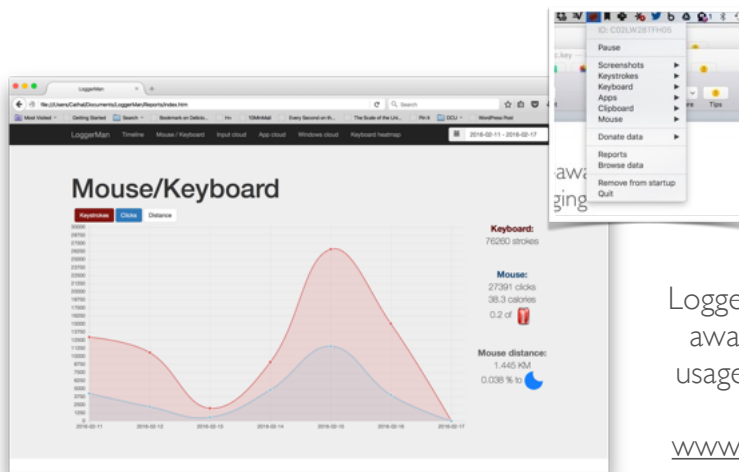
And other sources, such as location, music listening, BCI, etc...



1		<b>Gabriel Yared</b>	1,955 scrobbles
2		<b>MAMAMOO</b>	1,881
3		<b>Queen</b>	1,349
4		<b>Radiohead</b>	1,184
5		<b>Elton John</b>	1,082
6		<b>BLACKPINK</b>	971
7		<b>Hans Zimmer</b>	804
8		<b>Billy Joel</b>	795
9		<b>The Divine Comedy</b>	664



Cathal Gurrin (2026)



Loggerman: Privacy-aware computer usage data capture

[www.loggerman.org](http://www.loggerman.org)

Z. Hinbarji, R. Albatal, N. O'Connor and C. Gurrin (2016) *LoggerMan, a comprehensive logging and visualisation tool to capture computer usage*. In: 22st International Conference on MultiMedia Modelling (MMM 2016), 4-6 Jan, 2016, Miami, FL

Cathal Gurrin (2026)

## Aligned and integrated, this becomes a continuous lifelog...

```

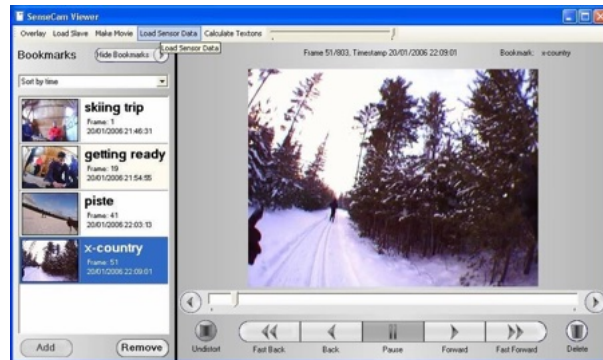
<NTCIR-lifelog-dataset>
<users>
  <user id="u1">
    <days>
      <day>
        <date>2016-08-15</date>
        <image-directory>NTCIR-Lifelog-dataset/u1/2016-08-15/</image-directory>
        <biometrics>
          <weight>81.89</weight>
          <fat-mass>19.78</fat-mass>
          <heart-rate>67</heart-rate>
          <systolic>125</systolic>
          <diastolic>84</diastolic>
        </biometrics>
        <activities>
          <steps>9340</steps>
          <distance>8464.468</distance>
          <elevation>26.87</elevation>
        </activities>
        <personal-logs>
        <health-logs>
          <time>06:10</time>
          <GLU>6.9</GLU>
          <BP>125/84</BP>
          <HR>67</HR>
          <mood>Anxious</mood>
          <comment>Diet change from today..</comment>
        </health-logs>
        <drink-logs>
          <log>
            <time>06:20</time>
            <drink>Coffee</drink>
          </log>
        </drink-logs>
        <food-logs>
          <log>
            <time>06:20</time>
            <food>Small fruit breakfast</food>
          </log>
        </food-logs>
      </day>
    </days>
  </user id="u1">
</users>
</NTCIR-lifelog-dataset>
</log>
</food-logs>
</personal-logs>
<minutes>
  <minute id="0">
    <location>
      <name>Home</name>
    </location>
    <activity>transport</activity>
    <bodymetrics>
      <calories>1.2</calories>
      <gsr>0.000183189</gsr>
      <heart-rate>59</heart-rate>
      <skin-temp>86.9</skin-temp>
      <steps>0</steps>
    </bodymetrics>
    <images>
      <image>
        <image-id>u1_2016-08-15_052242_1</image-id>
        <image-path>u1/2016-08-15/20160815_052242_000.jpg</image-path>
      </image>
    </images>
    <phone-images>
      <phone-image>
        <phone-image-id>u1_2016-08-15_06.23.14_phone_1</phone-image-id>
        <phone-image-path>u1/2016-08-15/phone/2016-08-15_06.23.14.jpg</phone-image-path>
      </phone-image>
    </phone-images>
    <music>
      <song>Runaway Train</song>
      <song-mbid>bb7df441-35b2-4c0d-99b5-6560a5bbbc51</song-mbid>
      <artist>Soul Asylum</artist>
      <artist-mbid>b10db9ad-b4c3-47f3-a7a4-37864b134f65</artist-mbid>
      <album>Black Gold: The Best Of Soul Asylum</album>
      <album-mbid>4f2ff67a-d196-48a6-ba0a-bff6724b94ec</album-mbid>
    </music>
  </minute>
</minutes>
</day>
</days>
</user id="u1">
</users>
</NTCIR-lifelog-dataset>

```

Cathal Gurrin (2026)

Example Lifelog systems from the early years (pre-2015)

Manual Event  
Segmentation in the  
Original Microsoft  
Sensecam Browser in  
2002

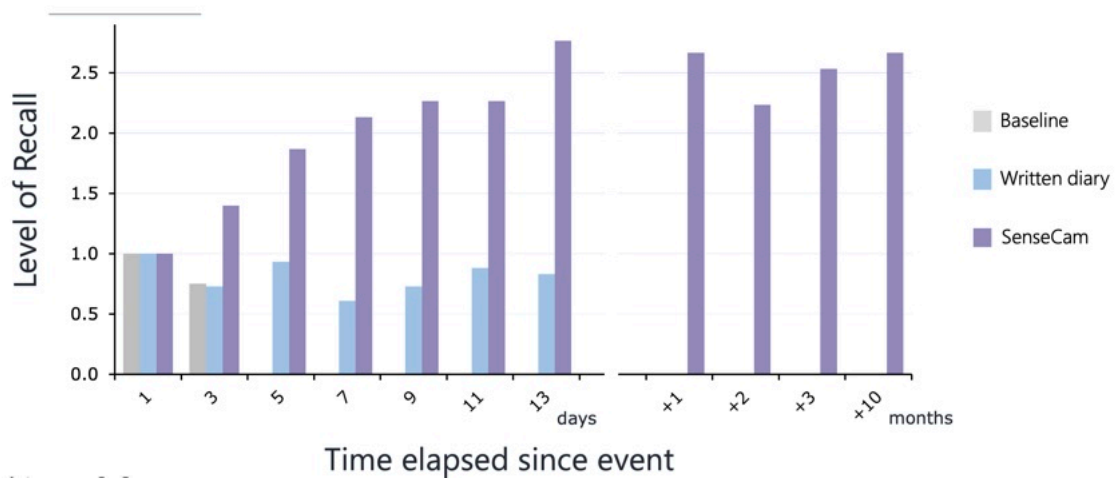


S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N Kapur, K. Wood, P. Dourish, A. Friday. *SenseCam: A Retrospective Memory Aid*. UbiComp 2006

Cathal Gurrin (2026)

Courtesy of Prof Steve Hodges (Lancaster University), creator of SenseCam

SenseCam clinical study of amnesia patient



Cathal Gurrin (2026)

# Event Segmentation



Events are analogous to our episodic memory and can easily be segmented using heuristics or basic models

A Doherty, C Ó Conaire, M Blighe, A.F. Smeaton, N.E. O'Connor. Combining image descriptors to effectively retrieve events from visual lifelogs. In: MIR 2008 - ACM International Conference on Multimedia Information Retrieval, 30-31 October, Vancouver, Canada.

Cathal Gurrin (2026)

# Public Health Studies

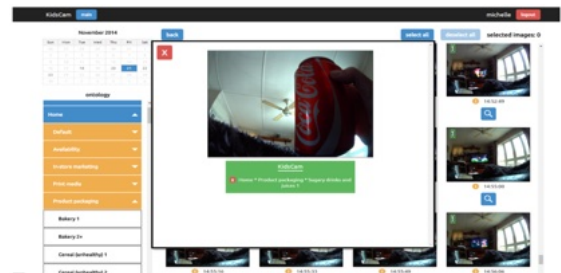
The Kidscam food/diet analytics system that we developed for U. Otago in New Zealand examines the daily life of 200 schoolchildren to understand their exposure to unhealthy food advertising.

It has been replicated in Beijing, China (U. Peking, Dept of Public Health) and extended to explore:

**Food marketing** - of non-core (fast food) advertising was seen over 27 times per day.

**Alcohol / Smoking** - Children are exposed to alcohol marketing on 85% of their visits to supermarkets.

Also, considering **Dangerous Environments** and **Cooking Activities**



T.Chambers, A.L.Pearson, J.Stanley, M.Smitha, M.Barra, C.Ni Mhurchucu, L.Signal. *Children's exposure to alcohol marketing within supermarkets: An objective analysis using GPS technology and wearable cameras.* In: *Health & Place*, Volume 46, July 2017, Pages 274-280.

Zhou Q., Wang D., Mhurchu C.N., Gurrin C., Zhou J., Cheng Y. & Wang H., *The use of wearable cameras in assessing children's dietary intake and behaviours in China*, *Appetite* (2019), doi: <https://doi.org/10.1016/j.appet.2019.03.032>.

Cathal Gurrin (2026)

Courtesy of Dr  
Allie Tran (DCU)



An objective assessment of children's exposure to brand marketing in New Zealand (Kids'Cam): a cross-sectional study  
Leah Watkins <sup>1</sup>, Ryan Gage <sup>2</sup>, Moira Smith <sup>3</sup>, Christina Mc Kerchar <sup>4</sup>, Robert Aitken <sup>1</sup>, Louise Signal <sup>3</sup>

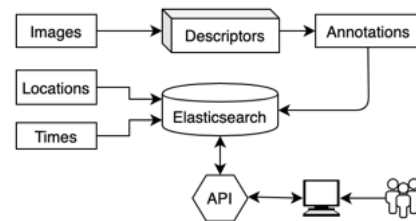
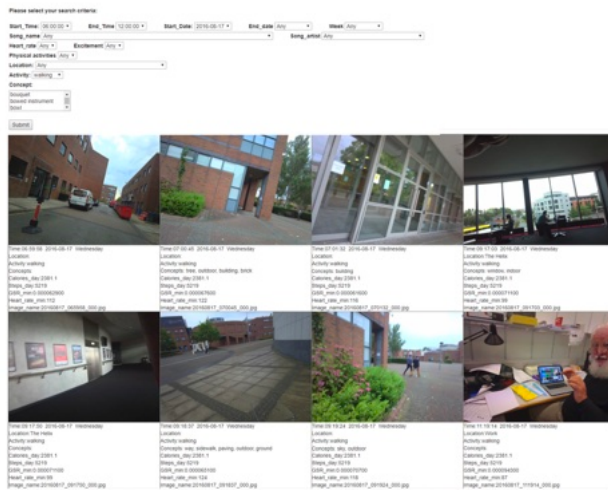


Courtesy of Dr  
Allie Tran (DCU)

Smith, M., Chambers, T., Abbott, M., & Signal, L. (2020). High stakes: Children's exposure to gambling and gambling marketing using wearable cameras. *International Journal of Mental Health and Addiction*, 18(4), 1025-1047.



# Which Facilitated Search Engines like this



L. Zhou, Z. Hinbarji, DT Dang-Nguyen & C. Gurrin. 2018. *LIFER: An Interactive Lifelog Retrieval System*. In *Proceedings of the 2018 ACM Workshop on The Lifelog Search Challenge (LSC '18)*. Association for Computing Machinery, New York, NY, USA, 9–14. <https://doi.org/10.1145/3210539.3210542>



## State-of-the-art from Grand Challenges

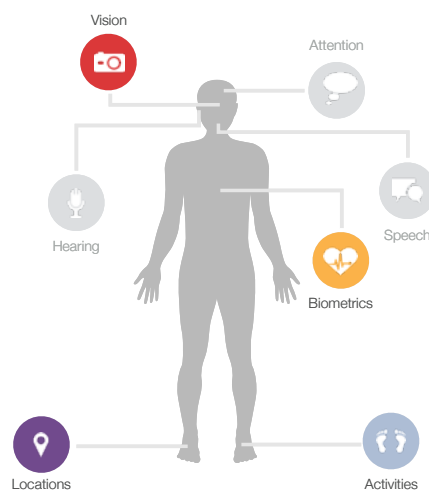
It was 2015, and there were no datasets, no evaluation measures or criteria, not even an understanding of what we should be doing. So, we started some grand challenges for the community.

So the first task was to create datasets, but who is willing to share their life data?

How can we release this dataset and maintain privacy of the wearer and bystanders?

Lifelogs are fundamentally different from standard multimedia datasets because they are:

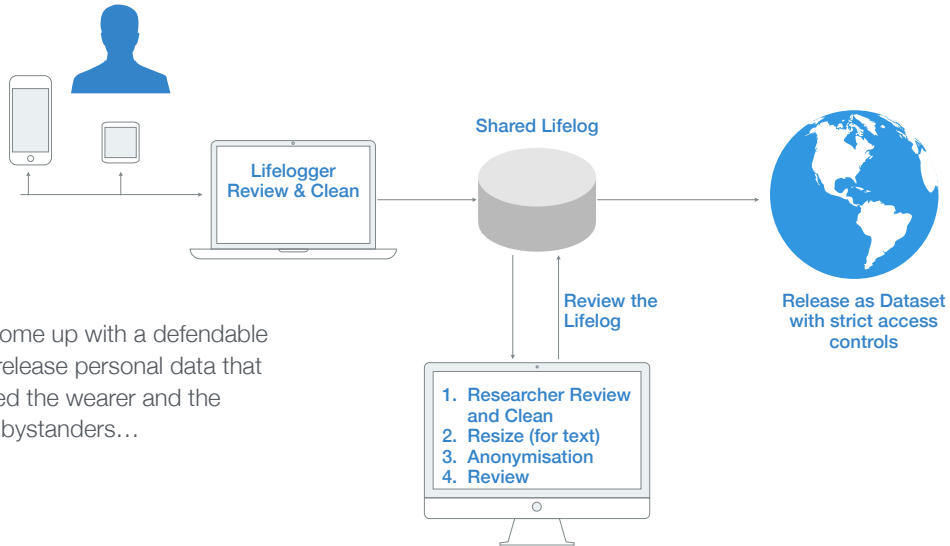
Uncommon  
Continuous (multi-year)  
Temporally Structured  
Highly Personal



They can contain:

egocentric video/  
images,  
audio,  
biometrics,  
locations,  
physical and  
information access  
activities, active  
annotations,  
... and various other  
passive and active  
sources

# A New Data Release Process



We had to come up with a defensible process to release personal data that protected the wearer and the bystanders...

Cathal Gurrin (2026)



So, we had a dataset, now we needed an evaluation system (DRES) and an understanding of the types of information needs that we needed to solve.

The screenshot displays the DRES interface for a competition run. At the top, the system name 'DRES' and a timer '0:045ms' are visible. The navigation bar includes 'Competition Builder', 'Competition Runs', 'User Management', 'Media Collection Management', and 'Audit Logs'. The user 'cathal' is logged in.

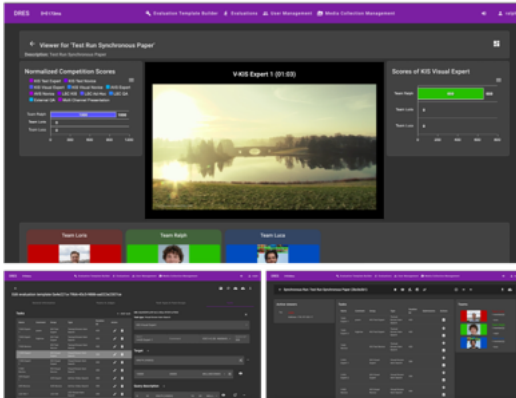
The main content area is titled 'Viewer for 'Official Run'' and includes a description: 'Description: REAL LSC2021 Competitors! This is not a test! :)'

There are three main panels:

- Normalized Competition Scores:** A horizontal bar chart showing scores for various competitors. MyScaal has the highest score at 100, followed by SomHunter (95), CVHunter (90), Voxento (85), and LifeSeeker (84).
- Task 25 (01:09):** A central text box containing the task description: 'Birds in a cage, a yellow one on the lower left. There was also one box with a small, GREEN old car (Beetle-like). No, the car was BLUE! It was in 2018 in May, I think it was a sunday.'
- Scores of LSC:** A horizontal bar chart showing scores for various LSC (Life Search Challenge) categories. Voxento has the highest score at 95, followed by SomHunter (90), CVHunter (85), and LifeSeeker (84).

At the bottom, a grid of competitor cards is displayed, each showing the competitor's logo, name, and score. The scores are: Exquisitor (0), LifeGraph (0), vitriv-VR (0), vitivr (0), lifeXplore (62), VIRMA (0), MyScaal (95), LifeSeeker (68), Memento (95), NTU-ILRS (94), PhotoCube (0), LifeMon (91), XQC (0), Voxento (96), SomHunter+ (95), FIRST (0), and CVHunter (90). Each card also features a small icon representing the competitor's performance.

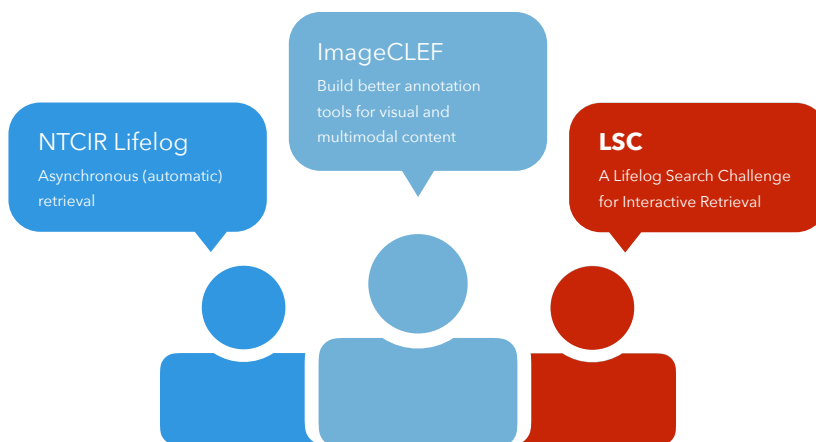
## Distributed Retrieval Evaluation Server (DRES)



- Flexible architecture supporting broad range of retrieval tasks
- Supports local and distributed settings synchronously and asynchronously
- Powers the Lifelog Search Challenge and the Video Browser Showdown
- Free and Open Source, available from <https://dres.dev>

Cathal Gurrin (2026)

## Three International Challenges



Cathal Gurrin (2026)

# Lifelog Search Challenge @ ICMR



Cathal Gurrin (2026)



What are the types of information needs that we could ask our personal lifelogs?

Query Type	Reason	Example	Category
<b>Known-Item Retrieval</b>	Finding a specific memory, object, place, or event that the user partially remembers.	"Find the photos from the restaurant in Singapore where I met the ACM Web steering committee."	Known-Item
<b>Behaviour &amp; Habit Analysis</b>	Understanding long-term patterns, routines, or behavioural changes.	"How has my drinking Guinness beer frequency changed over the past year?"	Ad-hoc
<b>Personal Insight &amp; Reflection</b>	Exploring emotional, social, or psychological aspects of life experiences.	"When do I appear most stressed during conference organisation periods?"	Q&A
<b>Decision Support &amp; Prediction</b>	Using lifelog evidence to guide choices or anticipate future outcomes.	"Does frequent international travel negatively affect my sleep and productivity?"	Q&A
<b>Episodic Reconstruction</b>	Rebuilding a sequence of events or experiences from a period of time.	"Reconstruct my trip to Cambodia during October 2025."	Recreation
<b>Prospective Memory &amp; Assistance</b>	Supporting future intentions, reminders, and unfinished tasks.	"Remind me to contact the researcher I met after the CASTLE workshop."	Reminder
<b>Identity &amp; Life Narrative</b>	Constructing higher-level stories, themes, and meaning from life experiences.	"Summarise the major stages of my research career in lifelogging."	Summarisation

KNOWN ITEM



Find any image of a particular event or activity described in an information need. E.g. find my keys.

AD HOC



Find as many images of something as you can, based on a textual information need. E.g. How many times did I go drinking?

Q & A



Answer a specific question. E.g. what is my favourite brand of Beer?

Cathal Gurrin (2026)

KNOWN ITEM

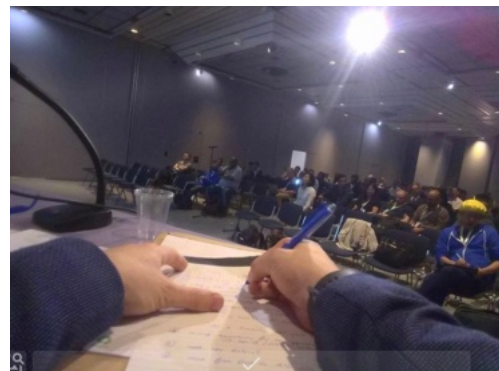


There was a man in the front row with a yellow hat on. **I was on a stage in a room with a lot of people watching.** I was on some sort of panel and writing notes on paper. **I remember the man had a blue sweater/top on also.** It was in France at ACM MM2019.

AD HOC



Q & A



Cathal Gurrin (2026)

# A Known Item Topic

0, 30, 60, 90, 120, 150... 300 seconds

I see Steve Wozniac...

I see Steve Wozniac on a wall of portraits.

I see Steve Wozniac on a wall of portraits. The wall was a brick wall with a door and large heater.

I see Steve Wozniac on a wall of portraits. The wall was a brick wall with a door and large heater. I was speaking to an audience before seeing the photos.

I see Steve Wozniac on a wall of portraits. The wall was a brick wall with a door and large heater. I was speaking to an audience before seeing the photos. I left by driving back to work.

I see Steve Wozniac on a wall of portraits. The wall was a brick wall with a door and large heater. I was speaking to an audience before seeing the photos. I left by driving back to work. It was in 2015 in March on a Wednesday.



Cathal Gurrin (2024)

KNOWN ITEM



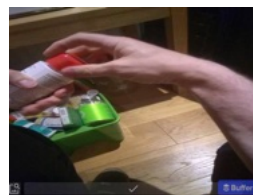
AD HOC



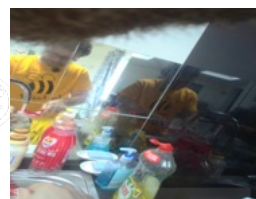
Q & A



Find examples of when I was taking medication or vitamins at home...



Find examples of when I was wearing my yellow 'Bee Happy' t-shirt



Cathal Gurrin (2026)

KNOWN ITEM



AD HOC



Q & A



I visited a famous Korean temple with friends a few years ago. What did I do immediately afterwards?

<Get a Coffee>



Damn it, my car has a flat tyre. What was the name of the car service/repair company that I used in the summer of 2019? I want to call them to get my car fixed.

<Tommy Tucker Tyres>



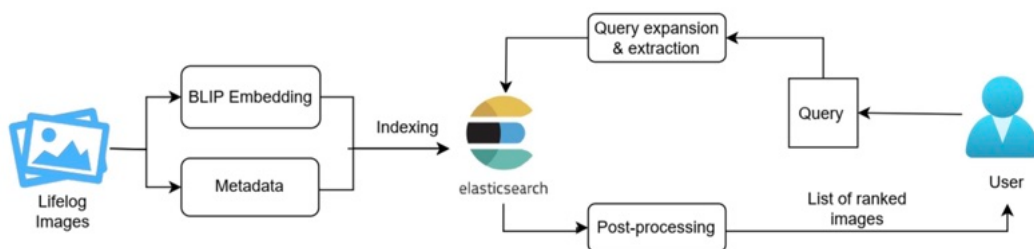
Which airline did I fly with most often in 2019?

<Turkish Airlines>



Cathal Gurrin (2026)

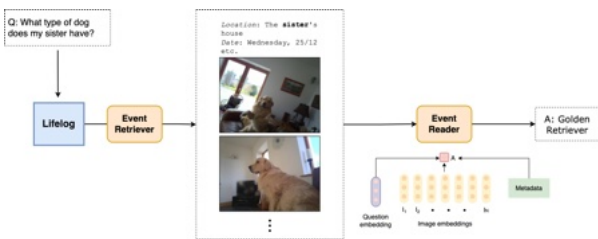
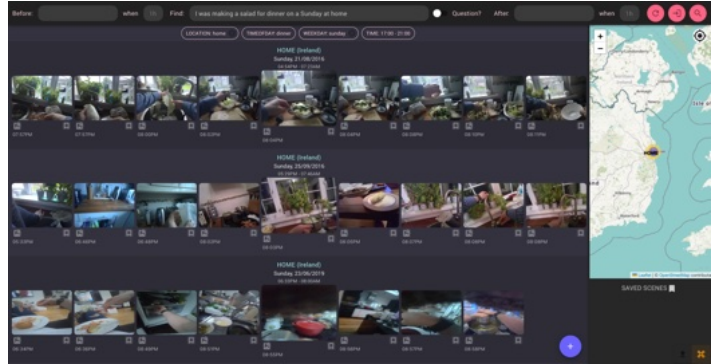
## What does a standard system look like?



Cathal Gurrin (2026)

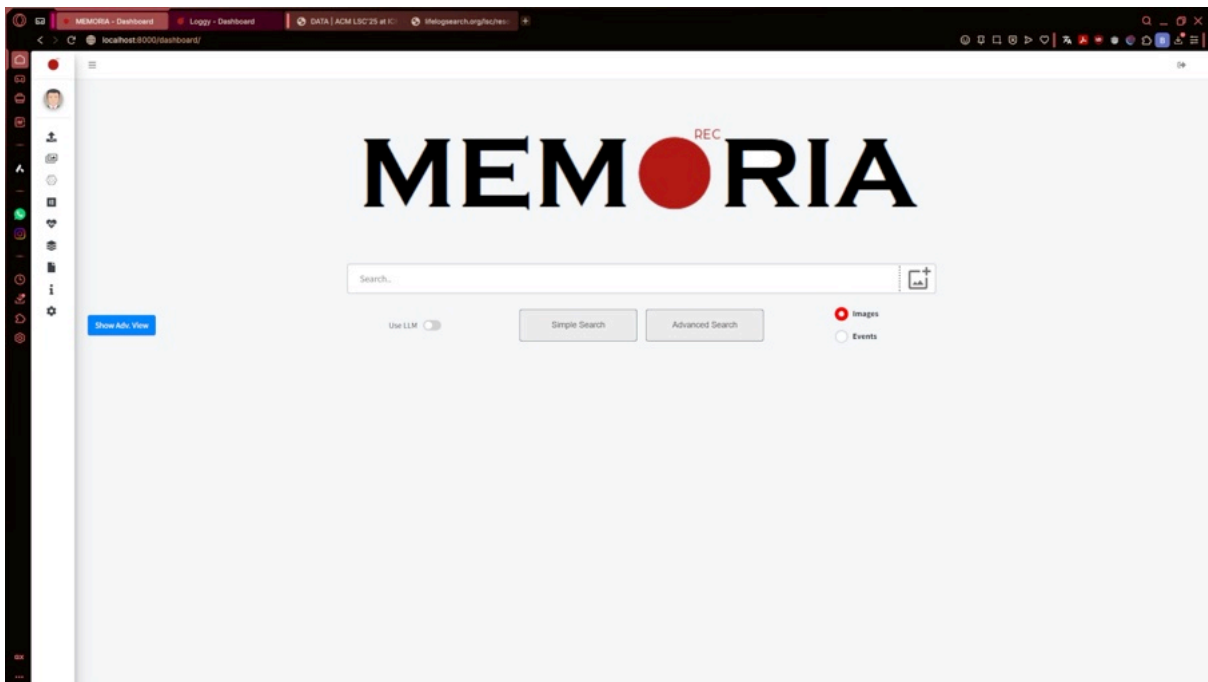
# MySceal/MyEachtra (2020-2024)

- Free-text search** (temporal search)
- Faceted search** via filters (automatic filter recognition)
- Embedding Models** for Visual Search
- LLM-based query interpretation**
- UI with Mapping**

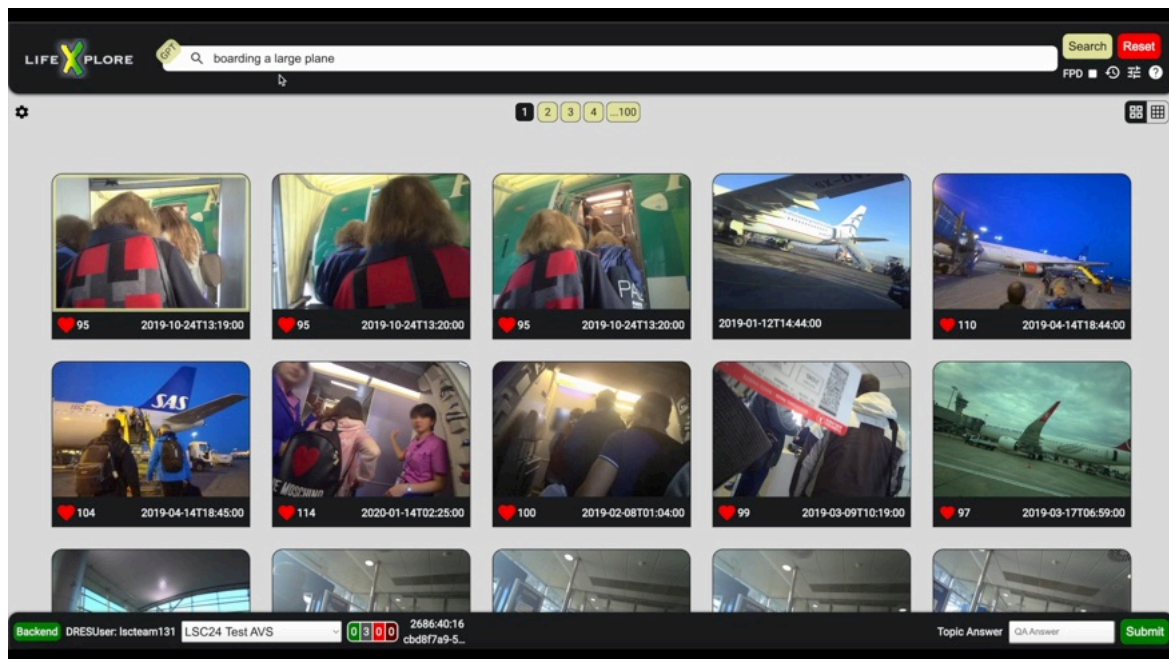


Ly Duyen Tran and Cathal Gurrin. 2025. *MyEachtraX: Semantic Event Retrieval with Time-Weighted Fusion*. In *Proceedings of the 8th Annual ACM Lifelog Search Challenge (Chicago, US) (LSC '25)*.

Cathal Gurrin (2026)



Alexandre Gago, Bernardo Kaluza, and António J. R. Neves. 2025. *MEMORIA: A Memory Enhancement and MOment Retreval Application at the LSC2025*. In *Proceedings of the 8th Annual ACM Lifelog Search Challenge (Chicago, US) (LSC '25)*.



Martin Rader and Klaus Schoeffmann. 2025. *lifeXplore 2025: Is Search in Text Generated by ChatGPT and BLIP2 better than in Embeddings from OpenCLIP?*. In *Proceedings of the 8th Annual ACM Lifelog Search Challenge (Chicago, US) (LSC '25)*



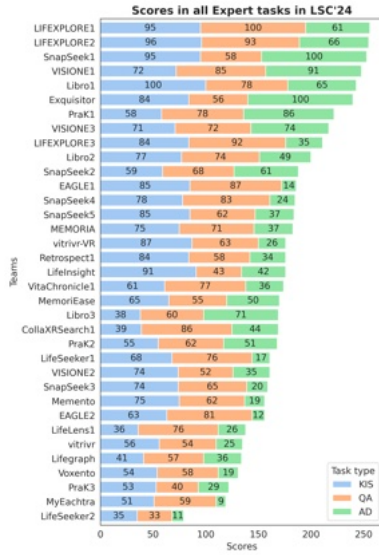
Nhut-Thanh Le-Hinh, Cong-Triet Huynh, Minh-Quan Ho-Le, Duy-Khang Ho, Minh-Triet Tran, and Viet-Tham Huynh. 2025. *LUMINA-1: Learning and Understanding Multimedia in Immersive Navigable Archives for Lifelog Retrieval*. In *Proceedings of the 8th Annual ACM Lifelog Search Challenge (Chicago, US) (LSC '25)*

## A lot of Variation in Approaches

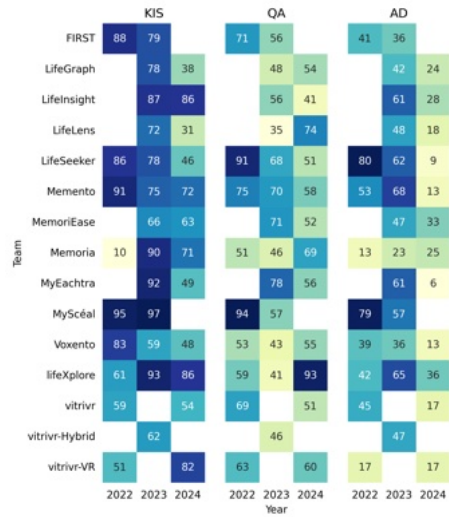
Index Methods	Text Handling	Image Handling	Metadata Handling	Query Types	User Interface	Devices	Other Issues
Conventional Text Inverted Index	aTFIDF / BM25	Visual Concepts (many approaches)	Manual Annotation	Auto Complete text	<b>Ranked Lists (images / events)</b>	<b>Desktop</b>	Interactive Learning (linear SVM)
NoSQL Database	OCR	Image Similarity	Metadata QE	Graph Querying	<b>Event Organisation of Results</b>	VR	Image quality enhancement
<b>Embeddings</b>	Query Expansion	Logo Detection	<b>Date/Time filters</b>	Localised Image Querying	SoMs	Mobile	Mapping data to Life Activities
Concept Hierarchy	W2WTF	<b>Visual Relevance Feedback / QE</b>	Semantically Enriched Metadata	<b>Temporal Queries</b>	Daily Summaries for Browsing	Voice + Screen	Single User Rec Sys.
<b>Event Segmentaiton</b>	PoS Tagging	External Examples for QE	Caregorisaiton / Clustering	Faceted Filtering	Multidimensional Space mapping		
BoWV				Map Search / Filter	Map Interfaces		
Knowledge Graphs				<b>Conversational Search</b>	VR Interfaces		
					3D interfaces		
					Visual Clustering		

So what have we learned from LSC 18-25?

We can quantify the performance of different systems by topic type

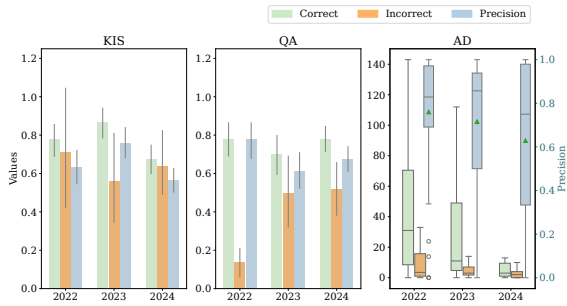


Compare system scores across multiple iteration) of the challenge



Cathal Gurrin (2026)

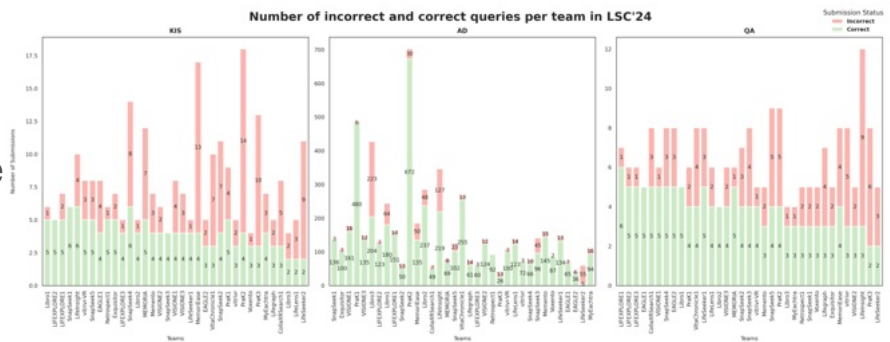
Correct and incorrect by task type and year



Exploring the difficulty of tasks over time.

And performance of team members since it is an interactive challenge

Number of incorrect and correct queries per team in LSC'24



Cathal Gurrin (2026)

# A Typical Interactive Lifelog Retrieval System

A typical lifelog retrieval system participating in LSC'26 is a sophisticated, AI-based platform that aims to support various retrieval tasks, including known-item search, ad-hoc search, and question answering, yet should be easy enough for a novice to use

Key characteristics include:

- **Multimodal Data Processing** and Indexing (vision LMs and data quality filtering)
- **Sophisticated Retrieval Strategies** (embeddings & hybrid models, event-based retrieval)
- **Conversational and Question-Answering** (QA) Capabilities (RAG & evidence-based answers)
- **Enhanced User Experience** (UX) and Interface (UI) Design (minimalist/adaptive UI, multiple device support, multiple types of result visualisation)

*Allie Tran, Duc-Tien Dang-Nguyen, Graham Healy, Steve Hodges, Bjorn Por Jonsson, Luca Rossetto, Klaus Schoeffmann, Minh-Triet Tran, Lucia Vadicamo, Cathal Gurrin. The State-of-the-Art in Lifelog Retrieval: A Review of Progress at the ACM Lifelog Search Challenge Workshop 2022-24. IEEE ACCESS 2025 (to appear)*

Cathal Gurrin (2026)

# 3

## Next Steps

Cathal Gurrin (2026)

These 2015-2025 challenges focused on visual data primarily and there was a problem....



How representative is the person who donated their lifelog data?



The released lifelog datasets typically have faces blurred and identifiable content removed (as best we could).

But think about your past? How many memorable events involve no-other-people?

It is a problem that faces should be blurred. Can we move to non-blurred datasets?



Cathal Gurrin (2026)

## Existing Datasets

Most of the current generation of datasets focus on:

- Action recognition
- Short clips
- Scripted activities
- Controlled environments

They lack:

- Long temporal continuity
- Real-world social interaction
- Multi-day behavioral patterns
- Fine-grained daily routines

**Ego4D** → the big egocentric benchmark (video, non-continuous, limited interactions, activity based, 933 participants)

**EgoRoutine** → routine modeling of activities (image, non-continuous, limited interactions, ADLs, 7 participants)

**LENA** → egocentric ADL video (non-continuous, ADLs, 10 participants)

**LSC datasets** → long-term but sparse egocentric (image, continuous multi-month or multi-year, ADLs, 1 participant)

**EgoLife** → egocentric and fixed ADLs for 1 week - 300 hours (video, audio, ADL, continuous, interactions, 6 participants, anonymized)

**CASTLE** → bring all these together into one coherent longitudinal video dataset of un-anonymised multimodal data from wearable & fixed cameras, 10 participants - 600 hours)

Cathal Gurrin (2026)



# CASTLE 2024 Workshop

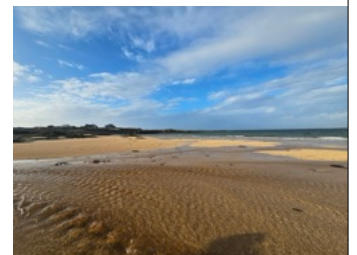
Supported by a SIG MM grant: One house. Four days in December 2024.  
Fifteen cameras. The workshop becomes the dataset.

The Parameters:

- **Camera-Free Zones.** Bedrooms and bathrooms remained private sanctuaries.
- **Full Review Rights.** Participants could review footage and request removal of any content (from all cameras)
- **No Anonymisation.** Everyone consented to be identifiable with full audio.
- **Data Governance.** Clear protocols for storage, access, and future use.
- **Minimal Third Parties.** Protected privacy of anyone outside the team

This dataset is unusual because we didn't anonymise anything. Instead, we built privacy into the process from day one.

Cathal Gurrin (2026)



Representing  
Ireland, Switzerland,  
Norway, Iceland,  
Austria, The  
Netherlands,  
Vietnam, and  
Thailand.

The team  
comprised 7  
professors, 2  
postdocs, and 3  
PhDs.



**Cathal Gurrin**  
Dublin City University



**Luca Rossetto**  
Dublin City University



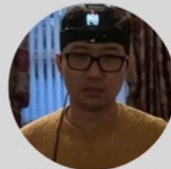
**Werner Bailer**  
JOANNEUM RESEARCH



**Klaus Schoeffmann**  
Klagenfurt University



**Allie (Ly-Duyen) Tran**  
Dublin City University



**Duc-Tien Dang-Nguyen**  
University of Bergen



**Björn Þór Jónsson**  
Reykjavik University



**Stevan Rudinac**  
University of Amsterdam



**Quang-Linh Tran**  
Dublin City University



**Hoang-Bao Le**  
Dublin City University



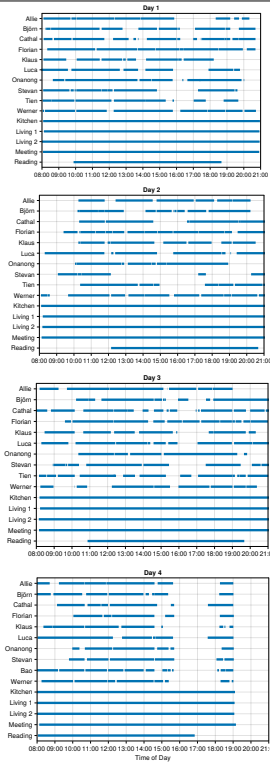
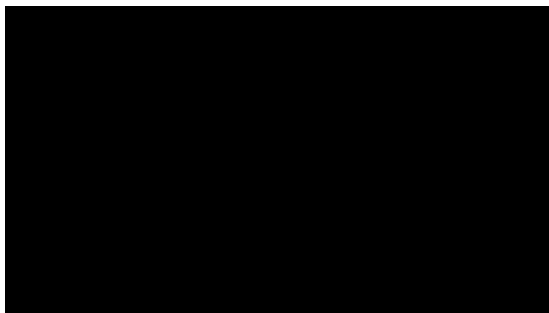
**Onanong Kongmeesub**  
Dublin City University



**Florian Spiess**  
University of Basel

# CASTLE 2024 Dataset

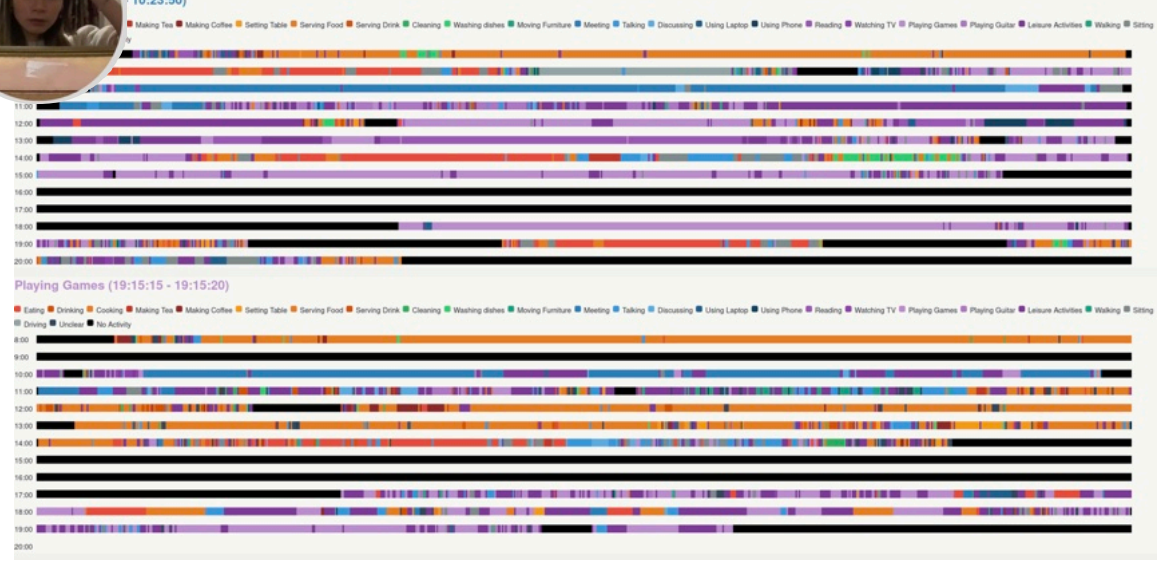
9 TB human activity dataset (600 hours from 10 people and 5 fixed cameras) in 4K, fully time aligned and non-anonymised  
 Free-living daily life activities with multiple daily interactions  
 Full audio transcripts and (soon) activity annotations



Luca Rossetto, Werner Bailer, Duc-Tien Dang-Nguyen, Graham Healy, Björn Þór Jónsson, Onanong Kongmeesub, Hoang-Bao Le, Stevan Rudinac, Klaus Schöffmann, Florian Spiess, Allie Tran, Minh-Triet Tran, Quang-Linh Tran, and Cathal Gurrin. 2025. *The CASTLE 2024 Dataset: Advancing the Art of Multimodal Understanding*. In Proceedings of the 33rd ACM International Conference on Multimedia (MM '25), Dublin, Ireland.



10:23:50)



Cathal Gurrin (2026)

# CASTLE 2024 Dataset



<https://castle-dataset.github.io/>

Cathal Gurrin (2026)

Courtesy of Prof Doc Tien Dang Nguyen (University of Bergen)

This wasn't a sterile lab experiment. We lived together for a week, documenting the full spectrum of mundane human behavior in naturalistic conditions.

## Cooking Together

Collaborative meal preparation with cross-cultural recipes and kitchen chaos

## Household Chores

The unglamorous reality of dishes, sweeping, and negotiating who does what

## Happy Quiz Night

Competitive trivia with dramatic reveals and questionable rule enforcement

## Reading Experiments

Structured tasks tracking attention, comprehension, and eye movement patterns

## Creative Activities

Painting sessions capturing collaborative artistic decision-making




## Card Games

Strategic gameplay revealing turn-taking dynamics and micro-negotiations

Courtesy of Prof Doc Tien Dang Nguyen (University of Bergen)

## CASTLE Challenge Tasks

We designed three benchmark tasks to test how well AI systems can navigate this multimodal haystack.

 <b>Event Search</b> <b>Goal:</b> Find specific moments in 600+ hours of video <i>Example query:</i> "When did someone spill coffee during the Happy Quiz?"	 <b>Object Search</b> <b>Goal:</b> Locate and track objects across multiple perspectives <i>Example query:</i> "Which room was the red notebook in on Wednesday afternoon?"	 <b>Video Question Answering</b> <b>Goal:</b> Answer natural language questions requiring cross-modal reasoning <i>Example query:</i> "Who was in the kitchen when the fire alarm went off?"
---	---	--

These tasks force AI systems to integrate video, audio, text, and sensor data, demonstrating the power and complexity of multimodal understanding.

Courtesy of Prof Doc Tien Dang Nguyen (University of Bergen)

## Examples of Hard Questions

Some queries require sophisticated reasoning across multiple data streams and temporal windows.

### "Who actually won the Happy Quiz?"

Requires: scorekeeping across multiple rounds, detecting disputes about rules, tracking point awards, identifying the final winner announcement

### "Who cleaned the kitchen voluntarily vs. reluctantly?"

Requires: identifying cleaning actions, analyzing facial expressions and body language, detecting verbal negotiations, inferring emotional state

### "What was everyone looking at during the painting activity?"

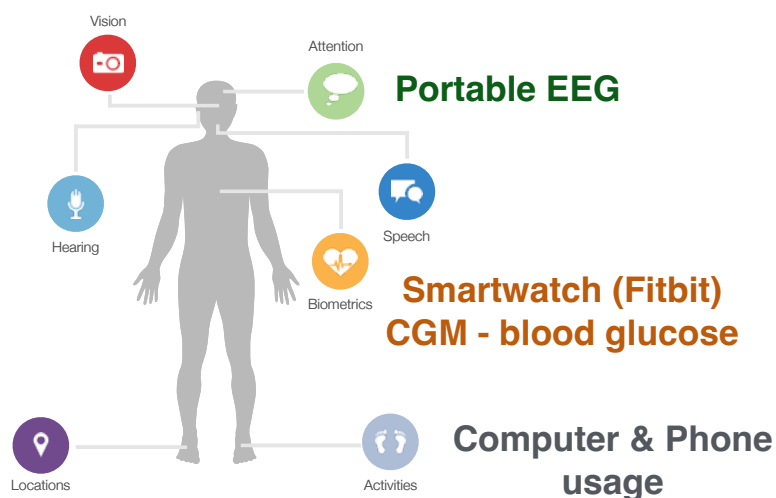
Requires: gaze tracking from first-person video, synchronizing multiple perspectives, spatial reasoning about scene composition

These aren't simple keyword searches. They demand contextual understanding, social reasoning, and synthesis across modalities—exactly the kind of challenges that push AI systems toward human-like comprehension.

## What is rich enough of a lifelog?

Smell, Taste, Touch, **Attention & Cognitive Load**, Emotional state, Stress, Mood, Momentary salience, Subjective enjoyment or boredom, Internal context.

Supported by another SIG MM grant: The same house. Six days in September 2025. Twenty cameras with similar parameters. This time we have more sensors, with an increased focus on biometrics.



Cathal Gurrin (2026)



# An Ideal Lifelog (individual & dataset)

- **Longitudinal** Duration (years not days, enough to train foundation models)
- Rich **Multimodal** Capture (synchronized)
- **Realistic** Everyday Behaviour (normal life, not staged)
- **Diversity** of Participants (age, culture, ethnicity, gender, lifestyle)
- **Human-centered** Evaluation Tasks with Rich Retrieval Ground Truth & Annotations
- Support for **Agentic AI** Systems
- **Privacy-Aware** Design (while maintaining usefulness and protecting bystanders and the individual)
- Not ignore the **Emotional and Social** Context
- **Ecological Validity** (normal life, lived normally)

Cathal Gurrin (2026)

76

But there are some societal challenges  
and risks here...

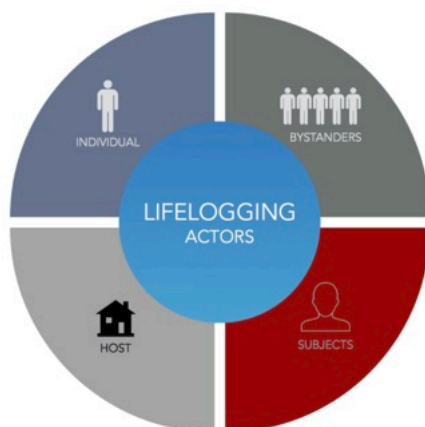
The same systems that can augment human memory and wellbeing can also enable unprecedented surveillance, manipulation, and loss of autonomy.

There are a number of key risks and challenges here...

Privacy	Exposure of intimate personal life
Surveillance	Misuse for monitoring and control
Psychological	Memory dependence and emotional harm
Social	Changes to trust and relationships
Technical	Security, hallucination, and misuse
Ethical & Legal	Ownership, consent, and human autonomy

## What about Privacy?

The meaning of privacy changes across different jurisdictions  
Different demographics have different expectations



- S. Ferdousa S. Chowdhury J. Jose. *Analysing privacy in visual lifelogging*. Pervasive and Mobile Computing, Volume 40, September 2017, Pages 430-449
- K. O'Hara, M. Tuffiel, N. Shadbolt. *Lifelogging: Privacy and empowerment with memories for life*. Identity in the Information Society, December 2008, Volume 1, Issue 1, pp 155-172
- A. Allen. *Dredging up the past: Lifelogging, Memory, and Surveillance*. The University of Chicago Law Review, Vol. 75, No. 1 (Winter, 2008), pp. 47-74
- R. Hoyle, R. Templeman, S. Armes, D. Anthony, D. Crandall, A. Kapadia. *Privacy behaviors of lifeloggers using wearable cameras*. UbiComp '14 - Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. Seattle, Washington — September 13 - 17, 2014
- R. Hoyle, R. Templeman, S. Armes, D. Anthony, D. Crandall, A. Kapadia. *Sensitive Lifelogs: A Privacy Analysis of Photos from Wearable Cameras*. CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. Pages 1645-1648. Seoul, Republic of Korea — April 18 - 23, 2015
- C. Gurrin, R. Albatal, H. Joho, K. Ishii. *A privacy by design approach to lifelogging*. In: O'Hara, K., Nguyen, C. and Haynes, P., (eds.) *Digital Enlightenment Yearbook 2014*. IOS Press, The Netherlands, pp. 49-73.

Cathal Gurrin (2025)

# Sousveillance

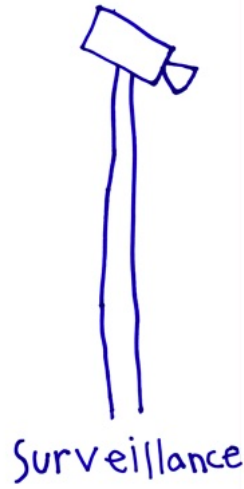
Steve Mann coined the concept of Sousveillance as a means of self-protection.

The watcher should be watched.

For periods of time he streamed live on the Internet.

S. Mann, J. Nolan and B. Wellman. *Sousveillance: Inventing and Using Wearable Computing Devices for Data Collection in Surveillance Environments*. In: *Surveillance and Society*, vol 1, pp 332-335. 2003.

Cathal Gurrin (2024)



Smart



Meta unveils two \$499 Ray-Ban smart glasses for prescription users

Cathal Gurrin (2026)



However, there are a lot of associated research opportunities.

The core challenges for our community can be summarised...

Theme	Core Challenge
Capture	Continuous sensing of human life
Understanding	Semantic interpretation of experiences
Memory	Long-term storage and temporal reasoning
Retrieval	Finding meaningful personal information
Personalisation	Adapting to unique individuals
Trust & Privacy	Secure, ethical handling of intimate data
<b>Agency</b>	<b>Proactive intelligent assistance</b>

In many ways, lifelogging can be one of the strongest real-world testbeds for future agentic AI systems.

## In Summary

I believe that lifelogs will become normal

Lots of opportunity for Agentic AI &  
Multimodal Systems to provide life-  
enhancing services

But there are risks and society might not be  
ready for this yet

# Thank You

## Cathal Gurrin

(@cathal - cathal@gmail.com)

