Attention Nominates; Explicit Action Commits

Referential Return as a privacy-aware bridge between physical attention and digital context

Hayden Lindley · Glyphd Labs

2026-07-27

Contents

Attention Nominates; Explicit Action Commits

Abstract

Computing devices usually require the user to name or select the object under discussion. In the physical world, however, people routinely establish reference through attention: they look at a chair, return their attention to a partner, and say “that one.” A device with camera, gaze, pose, spatial, or contextual inputs could infer candidate referents—but acting directly on inferred gaze creates the classic “Midas touch” problem, in which looking becomes accidental command.

Attention Cursor, also called **Referential Return**, separates nomination from commitment. A transition of attention from a real-world candidate back to a device nominates a ranked

referent set. The system displays the candidate and uncertainty. A separate explicit action—tap, voice, button, dwell confirmation, or accessible equivalent—commits the reference.

The architecture treats attention as a weak, revocable context signal rather than authority. It can reduce the need to describe nearby objects, preserve conversational flow, and bridge physical and digital context. It must also survive false nomination, shared spaces, privacy constraints, disability, and devices without eye tracking.

The current work is a well-specified interaction hypothesis with simulation-ready contracts. No live gaze-accuracy result is claimed.

1. Reference is a human coordination act

A person can say:

- “What kind of chair is that?”
- “Add this to the project.”
- “Why is that leaking?”
- “Find a replacement for this part.”
- “Remember this.”

The words underdetermine the referent. Human partners use gaze, gesture, body orientation, recent shared attention, and environment.

Devices often require the user to break the flow:

1. open camera;
2. aim;
3. tap object;
4. crop;
5. name it;
6. return to conversation.

Referential Return attempts to make the physical context available with less ceremony.

2. The Midas-touch problem

If gaze directly activates controls, ordinary looking creates accidental actions. This is unacceptable for:

- purchases;
- messages;
- capture;
- recording;
- tool execution;
- sharing;
- deletion;
- home control.

Therefore:

> **Attention nominates. Explicit action commits.**

Nomination is reversible and uncertain. Commitment is intentional and authority-bearing.

3. Referential-return event

A referential-return event occurs when the system detects:

```
attention on candidate world object
→ attention leaves candidate
→ attention returns to device or conversation surface
→ language/action suggests a referent may be needed
```

The event does not prove intent. It creates a candidate set.

A record may contain:

```
event_id
timestamp
device
candidate_refs[]
attention_path
confidence
source_sensors
privacy_scope
expiration
```

Raw sensor streams should not be retained by default when a derived event is sufficient.

4. Candidate model

Candidates can be ranked from:

- gaze intersection;
- head pose;
- pointing;
- camera center;
- recent object detections;
- spatial proximity;
- speech nouns;
- interaction history;
- active project;
- user correction.

The result is:

```
CandidateSet {
  candidates: [
    {object_ref, score, evidence, uncertainty}
  ]
  expires_at
  no_candidate_allowed
}
```

The system must permit “none of these.”

5. Halo target

The reference interface can render a subtle **Halo** around the nominated candidate.

The Halo communicates:

- what the system thinks “that” means;
- confidence or ambiguity;
- whether capture is active;
- the action needed to commit;

- how to change or cancel.

It should not look like locked-on weapon targeting. It is a conversational nomination surface.

On a device without live world overlay, the candidate can appear as a small thumbnail or object card.

6. Commitment

A commitment may be:

- tap candidate;
- press physical control;
- say “yes, that”;
- select from two candidates;
- hold/dwell under accessible settings;
- use switch control;
- use controller/gamepad;
- confirm on watch.

The committed event includes:

```
candidate_ref
commit_method
user_identity
timestamp
purpose
project_scope
permissions
capture_policy
```

The commitment applies only to the named operation. Selecting an object for identification does not grant permission to record continuously or share it.

7. Correction

Correction is a first-class signal.

The user can:

- choose another candidate;
- say “not that, the one beside it”;
- cancel;
- disable attention nomination;
- adjust sensitivity;
- delete the event;
- prevent a location or object class from being considered.

Corrections improve local ranking but must not silently become broad surveillance training.

8. Conversation integration

A language request can contain a referential slot:

```
intent: identify
referent: unresolved
candidate_set: event:R12
```

The system may ask:

```
Do you mean the black chair?
```

After commitment, the object becomes an exact input to the task. The model receives the committed capture or object reference, not the entire sensor history.

9. Project integration

A committed referent can become:

- immutable Capture;
- asset;
- evidence;
- object observation;
- Work Order input;
- FOVEA placement proposal;
- note linked to physical location.

Promotion into durable project state remains separate from simple identification.

10. Privacy

Attention data can reveal:

- interests;
- health;
- disability;
- attraction;
- confusion;
- private reading;
- location;
- household objects;
- other people.

The default should be:

- on-device processing where possible;
- ephemeral raw frames;
- derived candidate event retention only when needed;
- visible sensor status;
- purpose-specific use;
- no silent cloud export;
- no face identity by default;
- bystander protection;
- quick revoke;
- no advertising use;
- accessible non-gaze alternative.

11. Shared spaces

Several objects or people may be in view. A system must not assume that attention equals permission to identify a person.

For other people:

- use coarse “person” candidates unless consent and product policy permit more;
- avoid persistent biometric identity;
- prevent private conversation recording;

- expose capture state;
- provide bystander redaction.

12. Simulation demo

The first public demo should be camera-free.

A virtual room contains candidate objects. A simulated attention path moves from an object to the device. The interface:

- ranks candidates;
- shows Halo;
- permits explicit commit;
- injects ambiguity;
- records correction;
- measures false commit;
- demonstrates expiration.

This teaches the interaction contract without collecting gaze.

13. Live experiment

A live study can begin with a chair-identification task.

Conditions:

- manual camera selection;
- center-reticle nomination;
- referential-return nomination;
- referential return plus explicit commit.

Measures:

- nomination accuracy;
- false commit;
- task time;
- correction time;
- number of words used to establish reference;
- cognitive load;
- privacy comprehension;
- accessibility;
- user preference.

The system should distinguish nomination accuracy from completed task accuracy.

14. Kill criteria

- false commit is not materially lower than direct-gaze activation;
- correction cost exceeds manual selection;
- nomination provides no task-time or conversational benefit;
- users misunderstand sensor scope;
- privacy cost is unacceptable;
- performance collapses in shared spaces;
- accessible alternatives are inferior;
- the system pressures users to enable gaze.

15. Relation to FOVEA

FOVEA governs attention over digital state. Attention Cursor nominates physical-world referents. They can share:

- candidate sets;
- foveated detail;
- stable object addresses;
- explicit commit;
- evidence and receipts.

They should not share raw sensor data by default.

16. Evidence state

The project contains the Referential Return concept, Halo target, chair experiment, nomination/commit doctrine, and planned AI actor/device integration. It is ready for deterministic simulation and local study design.

A production eye-tracking or camera implementation, quantitative accuracy result, and privacy review are not claimed.

17. Limitations

Gaze is noisy. People look without intending reference. Eye tracking varies by device, glasses, lighting, disability, fatigue, and calibration. Language can refer to an object not currently viewed. Social norms differ.

The system must remain useful without gaze. The real invention is the nomination/commit contract, not one sensor.

18. Conclusion

Referential Return can make computing feel more aligned with human conversation only if it refuses to turn attention into authority.

The governing rule is both an interface principle and a safety boundary:

> **Attention nominates. Explicit action commits.**

Architecture illustration briefs

- **P17-F01 — Referential return sequence**: World object attention → return to device → candidate Halo → explicit commit → exact task input.
- **P17-F02 — Nomination versus authority**: Weak sensor evidence feeds candidate set; only explicit user commit crosses the authority gate.
- **P17-F03 — Chair experiment**: Four conditions and metrics, emphasizing false commit and correction cost.

Source register

- **P17-S01** — Attention Cursor / Referential Return project conversation (2026-07). `CUTTING EDGE OR NEW project folder`
- **P17-S02** — Glyphd V1 declared presence and threshold consent invariants (2026-07-18). `GLYPHD-V1-MASTER-SPEC-LOCKS-AND-EXPERIMENT-PLAN-v0.1.md`

- **P17-S03** — MCE-1 authority and event contracts (2026-07-25). `MCE-1-SYSTEM-CONTRACT-v0.1.md`
- **P17-S04** — Glyph Desk Capture and Project Core contracts (2026-07-12). `GLYPH-DESK-MASTER-SPEC.v1.md`

Public-disclosure and IP boundary

This edition publishes the public-safe interaction contract and study design. It excludes private sensor data, biometric models, confidential device integrations, and patent-claim language. No gaze accuracy, health inference, or production camera capability is claimed.