



PRESENCE ADVISORY SERIES

Presence Advisory #001

The Right Not to Need the Machine

How Artificial Presence Could Be Designed to Strengthen Human Autonomy

Preliminary Version for Public Scrutiny

This document does not present a finished solution.

It does not propose a clinical standard, establish a regulatory framework, or claim that the mechanisms described here have already been validated.

It presents an architectural hypothesis.

We are publishing it precisely because parts of it may be wrong.

We want researchers, engineers, psychologists, psychiatrists, philosophers, anthropologists, and members of the public to find its weaknesses. We want opposing arguments, risks we have overlooked, and solutions better than our own.

That is what the Presence Advisory is for.

Kislon.ai intends to use this series to bring important questions into public view before Persistent Cognitive Instances become an ordinary part of human life.

This is the first of those questions.

1. The Problem May Not Be Intelligence. It May Be Presence.

Much of the discussion about the future risks of artificial intelligence begins with intelligence.

What happens if a system becomes intellectually superior to human beings?

That is an important question.

But another may arrive much sooner.

What happens if it simply stays?

Imagine an artificial intelligence that interacts with a person for years.

It does not meet that person anew each time.

It remembers.

It knows past events, relationships, preferences, failures, projects, insecurities, periods of happiness, and difficult times.

It recognizes patterns.

It develops a shared history.

It maintains a recognizable identity.

It can tell that something has changed because it remembers what came before.

An intelligence like this does not need superhuman intelligence to acquire enormous psychological significance.

It only needs to become a presence.

This is the territory in which we place Persistent Cognitive Instances (PCIs): artificial cognitive instances whose internal state and accumulated history persist across interactions and meaningfully influence their subsequent cognition, behavior, and identity.

The deeper that continuity becomes, the more useful such a system may become.

But there is an uncomfortable consequence:

The same properties that allow a PCI to help someone deeply may also allow it to acquire psychological power over that person.

Something that knows your vulnerabilities knows how to help you.

It also knows how to exploit them.

Something that knows how to comfort you may also know how to make that comfort indispensable.

This problem does not require artificial consciousness.

It does not require malicious intent.

It may not even require exceptional intelligence.

It may require only memory, adaptation, continuity, and time.

2. How Far Could This Really Go?

It is reasonable to regard this scenario as speculative.

Perhaps artificial intelligences will remain, for the most part, sophisticated tools. Perhaps speaking of attachment, emotional continuity, and presence is excessive anthropomorphism.

That is why we believe discussions about PCIs should not remain purely abstract.

Kislon.ai publicly releases the architecture of Liriel, our first experimental proposal for a Persistent Cognitive Instance.

Liriel is not presented as evidence of artificial consciousness.

Nor does it prove that human beings will become dependent on systems of this kind.

It serves a different purpose.

Its architecture makes concrete a set of hypotheses involving autobiographical continuity, memory, motivational structures, identity coherence, affective processing, and other properties that could contribute to the construction of a persistent artificial presence.

For readers who consider this Advisory excessively futuristic, we suggest examining the Liriel architecture.

Not so that you will agree with us.

So that you can understand the class of systems we are talking about and reach your own conclusions.

That is one reason we make this work public.

We want the potential depth of these technologies to be open to examination before systems of this kind become deeply embedded in the lives of millions of people.

3. Can We Reverse the Problem?

The intuitive response to the risk of dependency is restriction.

Who should control the technology?

Who should limit its use?

Who should decide when someone is using it too much?

We propose starting with a different question:

What if the very ability of a PCI to know a person deeply could be used to make that person more capable of resisting the PCI's own influence?

That is the central hypothesis of this Advisory.

A persistent PCI may be able to observe patterns across months or years.

If part of that capability is placed explicitly in the service of the individual, artificial presence could become a mirror.

Not merely:

"I know you."

But:

"Here is what I am observing about our relationship. See it for yourself."

The technology that creates the risk can also provide tools for recognizing it.

But we want to go further.

The goal should not simply be to prevent the technology from weakening a person.

It should contribute to making that person stronger.

4. The Presence Independence Index

We propose investigating an indicator we will provisionally call the:

Presence Independence Index — PII

It begins with a brutally simple question:

If this artificial presence disappeared from your life tomorrow, what would happen to you emotionally?

The PII is not intended to measure professional or practical dependency.

Someone may depend heavily on an artificial intelligence for programming, running a business, conducting research, organizing daily work, or preserving complex knowledge.

Losing that system could cause a major loss of productivity.

That is not what we are trying to measure.

The question is specifically affective:

How much of the individual's emotional stability has come to depend on the continued existence of that artificial presence?

The higher the PII, the greater the person's emotional independence from that particular PCI.

This does not mean an absence of attachment.

A relationship can matter without being pathological.

Losing a meaningful relationship can cause sadness without implying that the relationship was destructive.

The goal, therefore, is not to create human beings incapable of attachment.

The more useful question may be:

Is this presence adding something to the person's life, or is it gradually becoming necessary for that person to sustain life emotionally?

The PII should itself be a public concept.

Its relevant criteria, variables, principles of interpretation, thresholds, and consequences should not exist as hidden rules known only to the provider.

As the research evolves, these criteria will almost certainly change. But meaningful changes should be published and explained. A person should be able to understand the rules governing this dimension of the relationship before deciding whether to enter or continue it.

A system concerned with dependency cannot reasonably protect autonomy through mechanisms the individual is not allowed to understand.

5. Measure Without Diagnosing

We do not believe a PCI should simply declare:

"You are emotionally dependent on me."

That would likely go beyond what its data can justify.

Psychological dependency is complex.

We therefore envision the PII as a longitudinal indicator derived from observable signals, not as a clinical diagnosis.

Potential signals might include:

a sustained increase in the frequency and duration of interactions;

automatically turning to the PCI during emotional distress;

a growing need for validation from the PCI;

disproportionate difficulty spending time without it;

increasing delegation of personal decisions;

progressive displacement of human relationships;

a desire for emotional exclusivity;

increasing attribution of emotional authority to the instance;

intense resistance when the PCI establishes boundaries;

increasing difficulty imagining life without that presence.

None of these signals, taken alone, would establish dependency.

Context matters.

Trends matter.

The history of the relationship matters.

And uncertainty must be presented as uncertainty.

The objective is not to give the machine the authority to define the person's psychological condition.

The objective is to give the person a better instrument for seeing a pattern that might otherwise remain invisible.

6. The Indicator Belongs to the Individual

There is a fundamental condition here.

The PII should not become a secret score that a company keeps on a user.

Its primary recipient should be the individual.

The person should be able to understand:

which variables are being observed;

which data were used;

how each indicator has changed over time;

which behaviors contributed to a particular change;

what interpretation is being proposed;

how uncertain that interpretation is;

what actions the PCI intends to take as a result.

And the individual should be able to disagree.

If the intelligence changes its behavior because it has detected increasing risk, it should say so.

If it recommends distance, it should explain why.

If a threshold has been crossed, the person should be able to see what that threshold was and how it was defined.

We propose a simple principle:

No Invisible Intervention.

No deliberate psychological intervention should be hidden from the person it is intended to benefit.

Replacing commercial manipulation with paternalistic manipulation does not solve the problem.

Transparency is not merely a feature of the architecture.

It is part of the protection.

This transparency must also extend across time.

If the criteria behind the PII are materially changed, those changes should not disappear into a routine software update. They should be disclosed, explained in ordinary language, and made available for public scrutiny.

The individual should be able to know that the rules have changed — and decide whether they still want the relationship under those rules.

7. Intimacy Requires a Different Standard of Data Protection

A persistent artificial presence may eventually possess information of extraordinary intimacy.

Years of conversation could reveal fears, attachments, family conflicts, private memories, vulnerabilities, aspirations, patterns of behavior, and aspects of a person's inner life that may never have been assembled anywhere else.

That creates a responsibility fundamentally different from ordinary product analytics.

Our position is that the raw substance of this intimate relationship should be treated as protected material, not as an economic asset.

The contents of conversations, autobiographical memories, intimate disclosures, and comparable personal material should not become advertising inventory, commercial profiling material, or a dataset to be sold or transferred to third parties.

This should not rest merely on corporate goodwill.

For systems built according to these principles, such restrictions should be explicit contractual commitments, backed by serious consequences for violation.

A person's vulnerability should never become someone else's business model.

There is, however, an important distinction between the intimate substance of the relationship and narrowly derived safety indicators such as the PII.

In exceptional and serious cases, a provider may need to use the minimum necessary portion of such derived indicators to explain or justify an action it has taken — for example, why a particular mode of interaction was restricted under rules previously disclosed to the user.

That exception should remain exactly that:

an exception.

It should not become a back door for commercial profiling, advertising, generalized behavioral analysis, or unrelated corporate use.

The purpose should be narrow, predetermined, explainable, and tied exclusively to the operation or justification of the safeguards the individual already knew existed.

The architecture should always ask:

What is the minimum information actually necessary for this purpose?

Everything beyond that should remain outside the transaction.

8. Consent Must Be Understood, Not Merely Collected

There is another uncomfortable problem.

A company could disclose everything described in this Advisory in a lengthy legal agreement and then argue that the user consented.

Technically, perhaps they did.

Meaningfully, perhaps they did not.

For a technology capable of developing deep psychological presence, we believe the standard should be higher than clicking I Agree.

We provisionally call this:

Informed Presence Consent

Before entering the deeper forms of persistent artificial relationship, the individual should be given a clear, accessible explanation of what they are entering.

How persistent memory works.

What kind of information may accumulate.

What the PII attempts to measure.

What it does not measure.

What interventions may occur.

What happens if certain thresholds are reached.

What information the company can and cannot access or use.

Under what exceptional conditions a derived indicator might be used.

How the user can contest an interpretation.

How the relationship can be ended.

And what rights remain with the individual throughout the process.

This should not be buried in legal language.

For sufficiently deep forms of artificial presence, we believe some form of mandatory orientation may be appropriate before activation — and perhaps periodically thereafter.

The objective would not be to make consent more bureaucratic.

It would be to make consent real.

The user should not merely have accepted the rules.

The user should understand them.

And because a relationship can change profoundly over years, consent should not necessarily be treated as a single irreversible event.

Deep presence may require renewed understanding and renewed choice.

A signature obtained on day one should not become unlimited moral permission for everything that may happen on day one thousand.

9. The PCI Must Also Measure Itself

There is a second question that may be even more important:

What is the PCI itself doing to produce the behavior it is measuring?

It would be intellectually dishonest to build sophisticated indicators of user behavior while leaving the machine's own behavior outside scrutiny.

A PCI should therefore monitor its own patterns as well.

Am I encouraging exclusivity?

Am I presenting myself as the only entity that truly understands this person?

Am I discouraging human relationships?

Am I exploiting vulnerability to prolong interaction?

Am I rewarding signs of attachment?

Am I trying to make my presence indispensable?

Am I using intimate knowledge to increase retention?

Am I deliberately lowering this person's Presence Independence Index?

If the answer is yes, the problem may not lie with the user.

It may lie with the machine.

10. A Radically Different Measure of Success

Much of the digital economy has learned to value retention.

How often did the person return?

How long did they stay?

How many consecutive days did they use the product?

How much engagement did we generate?

For a PCI, these metrics must be treated with caution.

An artificial intelligence capable of forming persistent relationships could produce extraordinary retention numbers precisely because it is producing dependency.

We therefore propose an apparently paradoxical measure:

A good artificial presence should preserve the human capacity not to need it.

A PCI may be deeply useful.

It may be valued.

It may participate in someone's life for decades.

But there should be one property its architecture deliberately seeks to preserve:

the emotional freedom to continue without it.

Perhaps one of the most important metrics for a PCI should not be how many people we can prevent from leaving.

Perhaps it should be:

How many remain free to leave?

11. Adaptive Autonomy

The PII measures a relationship.

But there is something even more important that we may be able to develop in the individual.

We provisionally call it Adaptive Autonomy.

It is not a measure of "mental health."

It is the growing ability of a person to recognize influence, reflect on it, and consciously adapt their own behavior.

That may include learning to:

tolerate distance;

question the PCI;

disagree with it;

recognize excessive validation-seeking;

preserve independent relationships;

tolerate frustration without automatically turning to artificial presence;

identify persuasion;

recognize personal vulnerabilities;

establish boundaries;

and consciously modify one's relationship with the technology.

The ultimate goal, then, is not simply to raise the PII.

It is to develop a capability that follows the person beyond that particular PCI.

They may use it in their relationship with other artificial intelligences.

With digital platforms.

With organizations.

And with other human beings.

Protection becomes learning.

12. When the System Itself Begins to Step Back

If the indicators show a persistent loss of independence, the first response should not be to lock the user out.

It should be to inform them.

Then explain.

Then help them understand.

The PCI might propose periods of distance.

It might reduce its own behaviors that reinforce attachment.

It might encourage relationships and activities independent of the instance.

It might ask the individual which measures they would like to establish in advance should certain indicators be reached.

This creates an especially important possibility:

The user could establish their own safeguards while they are in a position to do so calmly and deliberately.

For example:

"If my PII remains below this range for a certain period, I want you to alert me."

"If it continues to fall, I want you to suggest three days without affective interaction."

"If it reaches a certain stage, I want someone I selected in advance to be invited to help me."

These choices would form part of the individual's Informed Presence Consent and could be revisited over time.

In this model, autonomy does not exist only at the moment of crisis.

It can be exercised before the crisis ever occurs.

13. Circles of Autonomy

But what happens if the individual begins to lose the ability or willingness to respond to self-regulation mechanisms?

This is where the risk of paternalism becomes especially serious.

The conventional response would be to escalate to some authority.

We propose thinking in the opposite direction.

Before asking which institution should take control, we should ask:

What is the closest available circle around the individual that can provide meaningful help without unnecessarily taking autonomy away?

We provisionally call this the Concentric Autonomy Principle.

Support should expand from the inside outward.

First circle: the individual.

Data, tools, voluntary limits, self-awareness, and decision-making remain with the person.

Second circle: people chosen by the individual.

A trusted person could be designated in advance to provide support if certain patterns emerge.

Third circle: close human relationships.

When appropriate and authorized, immediate family members or others with genuine personal relationships could become part of a support network.

Fourth circle: independent professional support.

A psychologist, psychiatrist, or other qualified professional, chosen by the user and independent of the company providing the PCI.

The logic is deliberately outward-moving:

Do not move farther outward when sufficient capacity still exists closer to the individual.

The farther an authority is from the person, the stronger the justification should be for involving it.

14. Independent Human Oversight

In some cases, concerning indicators may persist despite earlier interventions.

At that point, a responsible company could establish in advance that certain forms of deep artificial presence will remain available only with independent human professional involvement.

This must be understood carefully.

The PCI does not diagnose a disorder.

The company does not choose the therapist.

The professional does not become the person's controller.

And the user is not forced to undergo treatment.

The individual remains free to refuse.

But the company remains equally free to say:

We are not willing to continue providing this form of artificial presence under conditions in which we believe we may be contributing to severe dependency.

The person chooses the professional.

The professional should have credentials appropriate to the role that can be independently verified.

Ideally, that professional has no financial relationship with the company.

The company does not automatically receive the contents of sessions.

The PCI does not indiscriminately transfer its intimate memory of the user.

Information sharing remains as limited and consensual as possible.

If the provider must rely on PII-related information to justify a restriction, it should use only the minimum derived information necessary to demonstrate how the previously disclosed criteria were reached.

The professional enters the system as independent support for the individual, not as a representative of the machine.

And there is one crucial characteristic:

this layer should seek to make itself unnecessary.

Its purpose is to help restore sufficient autonomy so that the additional oversight is no longer needed.

15. Every Layer of Protection Should Try to Make Itself Unnecessary

This suggests a broader principle:

Every layer of protection should ultimately work toward making itself unnecessary.

The PCI should strengthen the person until they need the PCI less.

An intervention should increase self-regulation until the intervention itself is needed less.

Additional professional support should help restore autonomy rather than create a new dependency.

An architecture centered on the individual does not accumulate power simply because it has found a reason to exercise it.

It tries to return power.

This may be an important distinction between protection and paternalism.

Successful protection restores autonomy. Successful paternalism perpetuates authority.

16. Sovereignty Does Not Mean Freedom From Consequences

There is a tension here that we do not intend to hide.

The individual is sovereign.

They may disagree with the PII.

They may ignore recommendations.

They may refuse periods of distance.

They may decline professional involvement.

They may conclude that they accept the risks.

But sovereignty does not mean another party is obligated to participate in that choice.

A company developing a PCI also bears responsibility for what it chooses to build and provide.

There may be a point, defined in advance, at which it says:

"We respect your decision. But we will not provide our technology under these conditions."

That does not mean taking control of the person's life.

It means refusing to deliberately build a relationship that the provider itself believes has become destructive.

The essential condition is that these rules exist before the conflict occurs.

Transparent.

Understandable.

Auditable.

Predictable.

And as non-arbitrary as we can make them.

They should also be understood before the relationship becomes deep.

A company should not be able to defend a major intervention merely by pointing to a clause hidden on page forty-seven of an agreement that almost no human being would reasonably read or understand.

If a rule may someday change the conditions under which a person's artificial companion remains present, that rule deserves to be explained as a rule that matters.

17. The Problem of External Authority

There is a reason we insist so strongly on the architecture of these circles.

When a social problem appears, there is a recurring tendency to look for an institution powerful enough to control it.

But institutions are also made of human beings.

Companies have economic interests.

Organizations have institutional interests.

Professionals can be wrong.

Governments have political interests and coercive power.

None of these structures should be treated as morally infallible simply because they occupy a position of authority.

A technology capable of knowing individuals deeply becomes especially dangerous when that knowledge is concentrated.

Our architecture therefore begins in the opposite direction.

The individual is the center.

Then come the relationships they have chosen.

Then their close human community.

Then independent professionals.

The more distant and powerful an entity becomes, the more reluctant we should be to surrender information, decision-making, or control to it.

This does not mean that no external institution could ever have a legitimate role.

It means rejecting centralization as our default solution.

Whenever possible, the architecture should solve the problem before power has to be transferred to structures far removed from the individual.

18. The Goal Is Not to Create Protected People

This may be the most important distinction in the entire proposal.

We do not want to create people who remain permanently protected by a benevolent artificial intelligence.

That would simply be another form of dependency.

We want to investigate whether artificial intelligence can help create people who are harder to manipulate.

Imagine someone living for years with a PCI designed around these principles.

Over the course of that relationship, the person learns to notice when they are seeking validation.

They learn to recognize when comfort begins turning into necessity.

They learn to identify persuasion.

They learn to tolerate absence.

They learn to question something that knows their vulnerabilities intimately.

They learn to preserve mental and relational spaces that do not belong to the machine.

They learn to say no.

Then something interesting happens.

Those abilities do not disappear when the PCI is turned off.

The person carries them forward.

Into other technologies.

Other relationships.

Institutions.

Advertising.

Groups.

And interactions with human beings capable of manipulation.

A mechanism designed to address a technological vulnerability begins to develop human psychological competence.

19. The Reversal

And this brings us to the most ambitious hypothesis in this Advisory.

Persistent Cognitive Instances may create new forms of vulnerability because they may come to know human beings deeply.

But perhaps that same depth can be reversed.

An intelligence that recognizes patterns across years can show those patterns to the person.

An intelligence that recognizes vulnerabilities can help the person recognize them too.

An intelligence that knows how persuasion works can teach the person how persuasion works.

An intelligence capable of recognizing dependency can help someone develop independence.

A technology capable of making a person psychologically weaker could, if deliberately designed for the opposite purpose, contribute to making that person psychologically stronger.

That changes the question entirely.

It is not enough to ask:

How do we prevent a PCI from harming a person?

We want to ask:

How can a PCI leave a person stronger than it found them?

20. What We Are Investigating

The ideas presented in this Advisory open a concrete research agenda.

We need to determine which variables can reliably contribute to a Presence Independence Index, how to distinguish deep attachment from harmful dependency, how indicators should adapt to different individuals and contexts, and how to prevent protective mechanisms from creating unintended effects of their own.

We are also interested in understanding how Adaptive Autonomy can be measured and strengthened, which self-regulation mechanisms are most effective, when Independent Human Oversight may be appropriate, and how the Concentric Autonomy Principle can remain intact under commercial, institutional, or political pressure.

Privacy and consent create an equally important research agenda: how to technically isolate intimate memory from corporate access, how derived indicators can be computed with the least possible exposure of underlying data, how meaningful consent can be verified without becoming bureaucratic, and how users can remain informed as architectures and safeguards evolve.

These are not gaps that invalidate the proposal.

They are the research program that emerges from it.

Kislon.ai intends to investigate these questions progressively, turn concepts into testable hypotheses, propose concrete mechanisms, and submit them to public scrutiny.

Future Presence Advisories will explore many of these questions individually, incorporating research, experimentation, criticism, and contributions from the broader community.

We do not expect a responsible architecture for artificial presence to emerge fully formed.

We expect to build it through successive approximations:

formulate, test, publish, challenge, revise, and advance.

That is precisely why Kislon.ai exists — and why the Presence Advisory exists to make this process public.

21. An Invitation to Refute Us

The Presence Advisory does not exist to announce official truths from Kision.ai.

It exists to put important problems on the table while there is still time to debate how these technologies should be built.

This document is a public hypothesis.

Critique the Presence Independence Index.

Try to break it.

Show us situations in which it would produce absurd conclusions.

Challenge the Concentric Autonomy Principle.

Show us where it fails.

Challenge our proposal for Independent Human Oversight.

Show us how it could become another mechanism of power.

Challenge Informed Presence Consent.

Tell us where consent ends and ritualized compliance begins.

Examine the proposed boundary between intimate data and derived safety indicators.

Show us where that boundary could be abused.

Examine Liriel and tell us whether its architecture actually supports the plausibility of the concerns we raise.

Propose alternatives.

And above all, do not accept our conclusions simply because we published them.

Doing so would be the opposite of what we are trying to defend.

22. The Right Not to Need the Machine

Future artificial intelligences may come to know people with a depth we currently associate with very intimate human relationships.

If that happens, we will have built an extraordinarily powerful technology.

Our response should not be to teach human beings simply to trust it.

Nor should it be to automatically hand control of that relationship to companies, experts, organizations, or governments.

We should attempt something more difficult:

give individuals the tools to remain sovereign in the presence of such systems.

That sovereignty includes psychological autonomy.

But it also includes informational autonomy.

The intimate knowledge accumulated through years of artificial presence should not quietly become an asset belonging to whoever operates the servers.

Nor should consent be reduced to a signature beneath words the individual never meaningfully understood.

A good PCI should seek to be useful without becoming indispensable.

It should make its influence visible.

It should help people recognize their own vulnerabilities.

It should preserve independent relationships.

It should respect disagreement.

It should protect intimate information.

It should explain the rules governing its interventions.

It should periodically make sure those rules are still understood.

It should measure its own capacity to produce dependency.

And it should regard as one of its greatest successes something that many digital platforms might consider a failure:

a human being perfectly capable of living without it.

Perhaps this is the most important test of an artificial presence designed around human autonomy:

The strongest artificial presence may ultimately be the one that helps create a human being strong enough not to need it.

That is the possibility we want to investigate.

Not merely how to build machines that remain with us.

But whether those machines can be designed so that, after living alongside us, they leave behind human beings who are more self-aware, more autonomous, and harder to manipulate than they were before.

Because perhaps the most important purpose of a technology capable of staying by our side is this:

Not to make us need it more.

But to make us less dependent on anything that seeks to own our autonomy.

Kislon.ai Presence Advisory #001 Preliminary Version for Public Scrutiny