# Cradle-to-Career Community Engagement Advisory Board Public Proposal Form

Instructions:

Per the Governance Manual proposal forms submitted will address significant gaps regarding feedback loops with data users, evidence-based decision making and analytical capacity, and equitable access to actionable information. Please note there should only be one proposal per form.

Name:

Nathan Pellegrin

Proposal Title: *no more than 50 characters*

Participant-Mediated Research Capacity

☑ By checking this box, I understand that if a Community Engagement Advisory Board member sponsors my proposal, it will be considered as a submission from the Community Engagement Advisory Board member.

Nature of Proposal:

☐ Feedback loops with data users
☑ Evidence-based decision-making and analytical capacity
☐ Equitable access to actionable information

What action should be taken to address the issue?

**Specific request**: The Board is asked to sponsor this proposal to initiate a structured feasibility, governance, and legal-design process — not to approve full implementation of a participant-mediated research (PMR) system at this time. The Office of Cradle-to-Career Data System should convene a time-limited workgroup to assess whether and how C2C could add a voluntary, participant-mediated research layer that would allow individuals to contribute additional information about their own lives and experiences, authorize selected data linkages, and express preferences about how their data may be used for approved research.

The proposed framework would draw on the model established by the NIH All of Us Research Program — which combines electronic health records with participant surveys, physical measurements, wearable data, and genomic data through informed consent and modular participant choice — while recognizing that C2C is a fundamentally different system, built as a linked administrative data infrastructure rather than as a research cohort. The workgroup's task is to determine what is feasible within California's specific legal and operational context.

**Why a workgroup, not immediate implementation**: The legal architectures of C2C and All of Us pull in opposite directions on the question of consent. All of Us is an opt-in research cohort governed by informed consent, a federal Certificate of Confidentiality, and a single central IRB. C2C operates under Education Code Sections 10860–10874 and an amendment to Civil Code Section 1798.24 (added by AB 132 in 2021) that creates a statutory carve-out permitting state agencies to disclose personal information to C2C without individual consent. Any PMR layer would therefore need to be grafted carefully onto a system whose foundational logic is administrative linkage rather than individual recruitment. This is not a barrier to participant-mediated research at C2C, but it does mean that the design, scope, and statutory authority of such a layer must be worked out deliberately before implementation.

**Workgroup charge — feasibility report deliverables:** The workgroup should produce a public feasibility report and implementation roadmap addressing the following questions:

1. Statutory authority. Whether C2C's existing authority under Education Code, together with the Civil Code disclosure pathway, is sufficient to support a voluntary, PMR layer, or whether new legislative authority is required. The report should specify what such legislation would need to include.
2. Legal and regulatory interaction. How a PMR layer would interact with the Family Educational Rights and Privacy Act (FERPA), the Health Insurance Portability and Accountability Act (HIPAA), the California Information Practices Act, the Confidentiality of Medical Information Act (CMIA), the Common Rule, and existing data-provider participation agreements. The report should identify regulatory or statutory gaps that would be necessary to support PMR.
3. Participant control. How participant consent, authorization, withdrawal, and ongoing preference management could function. The report should consider a modular consent design (similar in spirit to the All of Us Primary Consent, HIPAA Authorization, and tiered consent for return of results), accessible across languages and literacy levels, with comprehension checks and revocability.
4. IRB and human-subjects oversight. Whether human-subjects review should continue to be routed project-by-project through CPHS or whether a dedicated reliance arrangement or single IRB of record should be established for PMR studies.
5. Scope of participant-contributed data. Which data elements participants could voluntarily contribute — for example, surveys on educational goals, barriers to persistence, basic needs, caregiving responsibilities, advising experiences, mobility, job-search processes, and

detailed workforce experiences — and whether participants could authorize additional linkages to education, workforce, health, social-service, or employment records beyond what C2C already integrates.

6. Researcher access. How researchers would access PMR data, and whether a tiered access model (analogous to the All of Us Public, Registered, and Controlled tiers in the Researcher Workbench) would be appropriate for California.
7. Equitable design and governance. How communities historically underrepresented in research would be engaged in the design, governance, and oversight of the system. The All of Us experience is instructive: participants are recruited from groups historically underrepresented in biomedical research (Kozlowski et al., AJHG 2024), and research on missingness in that program (Cronin et al., PLOS ONE 2023) has shown that patterns of non-response are themselves informative about where trust-building must precede data collection.
8. Analytical justification. Why participant-contributed data is necessary, not merely useful. Administrative records document outcomes (enrollment, completion, employment) but cannot identify the unobserved variables — caregiving load, food and housing insecurity, advising quality, health events, immigration-related concerns — that confound causal inference about institutional effects. Within a counterfactual reasoning framework (Pearl & Mackenzie, The Book of Why, 2018; Ding & Li, Statistical Science, 2018), researcher-designed instruments administered to consenting participants are the standard mechanism for closing these gaps and distinguishing program effects from pre-existing differences among students.
9. Safeguards against harm. What technical, administrative, and policy safeguards are needed to prevent stigmatizing, discriminatory, or otherwise harmful uses of individual-level data, including protections against compelled disclosure, re-identification, and inappropriate secondary use.

10. Alternative architectures. Whether the goals of a participant-mediated research layer are best achieved by building inside C2C, by establishing a complementary California cohort hosted at a University of California medical or research institution and linked to C2C through narrow, consented bridges, or by some hybrid arrangement. The relative legal, operational, and trust implications of each option should be analyzed.

**Pilot recommendation:** The workgroup should consider scoping a small, low-risk pilot to test the operational and consent design — for example, a voluntary survey module for adult learners or transfer students, with explicit consent, transparent data-use rules, IRB review, and community oversight, linked only to approved research questions and evaluated before any expansion is considered.

Alignment with the CEAB charge: The proposal speaks directly to all three areas the Board is responsible for. It would create the structural feedback loop that C2C currently lacks: a mechanism for participants not only to contribute information about themselves but to receive research findings in return — a reciprocal model that All of Us has demonstrated at scale. It would expand analytical capacity by enabling researchers to move from describing disparities to explaining them. And it would advance equity by giving communities most represented in C2C's data a voice in the questions asked of that data and the protections that govern its use.

The immediate action requested is therefore modest in scope but consequential in direction: sponsorship of a feasibility, governance, and legal-design process that produces a public report and an implementation roadmap, positioning C2C and the Legislature to make an informed decision about whether and how to add this capacity.

Who could benefit from this action?

A participant-mediated research (PMR) layer would benefit a wide range of stakeholders, but the benefits are distinct enough that they should be considered separately rather than collectively.

Students and adult learners would benefit most directly, in two ways. First, PMR would give individuals an active rather than passive relationship to the data system that already represents them. C2C currently links administrative records about students; it does not currently offer students a structured way to add their own perspective on those records or to participate in shaping the research conducted from them. A PMR layer would create that channel.

Second, the model permits reciprocal information flow. For California students, the analogous benefit would be the ability to see how research findings about their cohort apply to their own pathway — turning C2C from a system that observes them into a system that informs them. Communities historically underrepresented in research stand to gain disproportionately if the system is designed well. Low-income students, students of color, adult learners with caregiving responsibilities, first-generation college students, undocumented students, students with disabilities, and others are heavily represented in C2C's administrative data but have no current mechanism to influence what is asked of that data or to contribute the context that explains the outcomes recorded. All of Us has demonstrated that participant-mediated systems can prioritize this engagement: more than 75% of its participants come from groups historically underrepresented in biomedical research (All of Us Research Program Investigators, "Data Quality, Utility, and Diversity," PMC9403360). The research on missingness further showed that patterns of non-response are themselves informative — they identify exactly where trust-building, community partnership, and culturally responsive engagement must precede data collection. A PMR layer at C2C, designed with community oversight from the outset, would create the same diagnostic capacity for California:

revealing where the system is and is not earning the trust of the communities it claims to serve, and providing a structural basis for repair.

Researchers and institutional research practitioners — including those at the community colleges, CSU, UC, the independent colleges, and policy research institutions — would gain analytical capacity that is not currently available from administrative records alone. C2C's data is excellent for documenting outcomes (enrollment, completion, transfer, earnings) but limited for explaining the mechanisms behind those outcomes. Within a counterfactual reasoning framework unobserved confounders — caregiving load, food and housing insecurity, health events, advising quality, immigration-related concerns, work schedules — make it difficult to distinguish the effect of institutional interventions from pre-existing differences among students. Australian education researchers have demonstrated this directly in the context of administrative-data policy research (arXiv:2301.06219, 2023): causal diagrams and valid variable selection critically depend on access to variables that administrative records do not capture. Researcher-designed instruments administered to consenting participants are the standard mechanism for closing these gaps.

Educators, advisors, and student-support practitioners would benefit because PMR-enabled research can produce findings specific enough to inform day-to-day practice. Administrative data can tell a counselor that transfer rates among Pell recipients lag by a measured margin; participant-contributed data can begin to indicate why — whether the binding constraint is advising quality, course availability, financial aid timing, family caregiving, or transportation. Findings of that texture are what move research from description to actionable guidance.

Community organizations and advocacy groups would benefit because PMR creates the structural feedback loops that the Community Engagement Advisory Board exists to support. The current C2C model permits community input into governance and dashboard design, but the underlying data flow remains one-directional: administrative records in, aggregated outputs out. A PMR layer would create a genuine bidirectional channel — community members can contribute, can shape what is asked, can see what is learned, and can hold the system accountable to its stated equity commitments. This is the feedback-loop architecture that distinguishes participatory data systems from purely administrative ones.

State agencies and policymakers would benefit from a markedly improved evidence base for program design and resource allocation. Equity analyses, intervention design, basic-needs program targeting, workforce planning, financial-aid policy, and evaluation of public investments all currently depend on data that documents disparities without fully explaining them. Adding participant-reported context to the administrative backbone would let agencies move from describing patterns of inequity to designing programs that address root causes — and would let the Legislature and Department of Finance evaluate the return on public investment with a richer signal than completion and earnings statistics alone provide.

Data providers — the agencies and segments that contribute to C2C — would benefit indirectly. PMR-enabled research that explains outcomes, rather than only describing them, gives data providers a stronger evidentiary basis for the policies they implement and the budget requests they make. It also creates a mechanism for surfacing data-quality issues from the population side: where participant-reported information systematically diverges from administrative records, that divergence points to gaps or errors in the underlying source data that providers can address.

C2C itself, finally, would benefit. The system's enabling statute frames it as a source of actionable data and research on education, economic, and health outcomes for individuals, families, and communities. A PMR layer would strengthen that mission by making the system not only about people but more directly informed by people.  The feasibility study proposed in the previous section is the path by which all of these benefits become accessible — but the benefits themselves are the reason to undertake it.

Who would implement this action?

Implementation can take more than one form, and the Community Engagement Advisory Board is well positioned to determine the approach that best fits its capacity and priorities.

One option is for the CEAB to convene a working group from among its own members — drawing in particular on those whose focal areas align most closely with the task, such as Research and Evaluation, Policy, and Community Engagement and Organizing. A member-led group could handle the early scoping work: defining the questions, identifying who should be consulted, and shaping the charge.

Another option is for the CEAB to recommend that the Office of Cradle-to-Career Data System lead the effort, in consultation with the Governing Board, the CEAB, the Data and Tools Advisory Board, data providers, legal and privacy experts, and representatives of communities whose data would be involved. This approach is suited to the phase of work that requires operational and legal capacity beyond what a Board working group can be expected to provide.

A combined approach is also possible, with the CEAB scoping and chartering the effort and the Office carrying it forward under that charter, with regular reporting back to the Board.

Whatever pathway is chosen, the work itself should be deliberate and inclusive. It should include public listening sessions, consultation with data providers, legal and privacy analysis, review of comparable models, consideration of consent and oversight structures, attention to community governance, and the option of a small pilot to test the design before any broader implementation. The process should be time-bounded, with interim updates to the CEAB and the Governing Board, and the final product should be a public report released for community input.

These recommendations are made with the understanding that close collaboration with data providers will be essential, since they hold authority over the underlying records and their continued participation is what makes the system valuable. If the work concludes that additional statutory authority is needed, the recommendation should be developed in partnership with the entities that would be affected.

How does the action relate to the mission and vision of C2C?

The proposal aligns with each of the three commitments expressed in C2C's mission and with the central themes of its vision, and it does so in a way that strengthens rather than departs from the system as currently constituted.

On the mission. C2C's mission is "to be California's source of actionable data and research on education, economic, and health outcomes for individuals, families, and communities" and "to expand access to tools and services to navigate the education to employment pipeline." A participant-mediated research (PMR) layer advances each phrase of that mission distinctly."

Actionable data and research" — The actionability of research findings depends on whether they can guide intervention. Findings drawn solely from administrative records can describe disparities (which groups complete, transfer, or earn at what rates) but typically cannot identify the mechanisms that produce them. Researcher-designed instruments administered to consenting participants are the standard mechanism for adding the variables — caregiving load, basic needs, advising experience, health events, work schedules — that turn descriptive findings into actionable ones. PMR is the architecture that lets C2C deliver on the "actionable" qualifier its mission already contains.

"Education, economic, and health outcomes." — The mission explicitly names health outcomes alongside education and economic outcomes. C2C currently integrates limited health-adjacent data through DHCS and DSS participation, but the system has no current mechanism for participants to contribute health-relevant context about themselves (chronic illness, disability, caregiving for a sick family member, mental-health barriers to persistence) in ways that can be linked to their educational and workforce trajectories. A PMR layer is the natural architecture for honoring the health-outcomes component of the mission without expanding the compulsory-disclosure footprint of the administrative system.

"For individuals, families, and communities." — The mission is addressed not only to researchers and policymakers but to the people represented in the data. PMR is the mechanism by which those individuals, families, and communities can do more than receive information from the system: they can shape what the system asks about them, and they can receive findings in return. This is consistent with the reciprocal model demonstrated by the All of Us Research Program (NEJM 2019), in which participants both contribute to and receive results from the research conducted on their data."

Expand access to tools and services." — PMR is itself an expansion of access: a new, voluntary tool through which Californians who wish to contribute more than their administrative record can do so, with transparent rules and revocable consent.

The vision states that C2C "connects individuals and organizations with trusted information and resources" and that the system "empowers individuals to reach their full potential and fosters evidence-based decision-making to help California build a more equitable future." Three phrases in particular bear on this proposal."

“Trusted information and resources." — Trust is built by structural features, not by assertion. A system in which agency data flows in and aggregated results flow out, has one trust architecture; a system that additionally offers individuals a voluntary, modular, revocable way to participate in research about themselves has a different and complementary one. The voluntary layer does not displace the administrative layer — it sits alongside it and gives Californians who want a more active relationship to the system a legitimate, governance-backed way to have one. This is a meaningful contribution to the trustworthiness of the system as a whole.

"Empowers individuals to reach their full potential." — The current empowerment offered by C2C is informational: better dashboards, better data stories, better planning tools. PMR adds a participatory dimension to that empowerment, allowing individuals not only to use the system's outputs but to influence its inputs and its research agenda.

"Evidence-based decision-making … for a more equitable future." — The equity commitment is the through-line of the vision, and it is also the strongest analytic case for PMR. Inequities in education and workforce outcomes are shaped by factors that are difficult to see in administrative records alone, and the people most affected by those inequities are also the people best positioned to identify which contextual variables matter. A PMR layer designed with community oversight allows the system to surface those variables in ways that are respectful, transparent, and analytically rigorous — moving equity work beyond description toward explanation and design.

On the apparent tension with "how we use data." The C2C website states that "the Office of Cradle-to-Career Data System does not collect new data from students or families; it links and enables the public to gain insights from data that has already been collected and validated by its data partners." This proposal does not contradict that statement; it proposes a deliberate, voluntary complement to it. The current model — linkage of data already collected — is the foundation of the system's scale, and nothing in this proposal disturbs that foundation. PMR would introduce a separate, voluntary stream in which consenting individuals choose to contribute additional information for research purposes, governed by informed consent, modular authorization, IRB oversight, and community governance. The two streams are architecturally distinct and legally distinguishable, and the feasibility study proposed in the prior responses is precisely the process by which that distinction would be designed.

In sum. The proposal extends the mission rather than redirecting it. It strengthens the actionability of the system's research, honors the often-underused health-outcomes component of the mission, deepens the system's relationship to the individuals it serves, and supports the vision's commitments to trust, empowerment, evidence-based decision-making, and equity. It does so without altering the existing administrative-linkage architecture, and it does so in a way that explicitly invites communities historically underrepresented in research into the design and governance of the new capacity. This is the participatory, ethical, and analytically powerful direction the mission and vision already point toward.