



Securing Tomorrow's Missions Today.



Before You Fund the Redesign: A Fixed-Price, Four-Week Usability Study of One Digital Service

User Testing & Feedback Analysis: real users attempting real tasks, the feedback the agency already holds, and a severity-rated backlog the program can act on

AvalonTechServices.com

contact@AvalonTechServices.com

Executive Summary	2
Why Services Fail Users Quietly	2
Why Teams Skip the Evidence	3
The Authorities and Where the Study Fits	4
What the Study Does	4
Tasks, Participants and Sessions	5
The Feedback Annex	5
Deliverables	6
Where the Study Stops	6
What the Evidence Can and Cannot Claim	7
Questions the Contracting Officer Will Ask	8
How a Study Runs	9
What the Agency Provides	9
What Is Not Included	10
How Avalon Staffs It	10
Pricing and How to Buy	11
Illustration: One Service, Eight Participants, Four Weeks	12
Labor Loading	12
What the Study Costs	13
What the Agency Brings to the Readout	13
Common Pitfalls	14
Conclusion and Next Step	14
Appendices	14
Appendix A: Glossary	14
Appendix B: Illustration Assumptions	15
Appendix C: References	16

Executive Summary

Most federal digital services are judged by the people who built them. The program office sees the requirements met, the help desk sees the calls, and leadership sees a completion rate that nobody can quite explain. What is usually missing is the simplest evidence of all: a record of real users attempting the tasks the service exists for, and where they fail.

That gap is expensive. A form that users cannot finish sends them to the phone line, the mail, or nowhere. A redesign funded without evidence often rebuilds the same problems in a new look. The 21st Century Integrated Digital Experience Act (2018) directs agencies toward user-centered websites and digital services, and the Paperwork Reduction Act shapes how agencies may collect information from the public. Neither is a reason to skip research; both are reasons to plan it properly.

This paper is for the program manager or product owner who owns the service and usually the operations and maintenance budget, the digital experience or web director and the CX lead who champion user-centered delivery, the CIO's digital services office, and the COR who buys the work. It describes Avalon Technology Services' User Testing & Feedback Analysis service, which starts with one engagement:

- **A fixed-price, four-week usability study of one digital service.** One unclassified service, one user segment, six to eight participants the agency supplies, moderated remote sessions on the agency's own Teams or Zoom for Government, and a feedback annex built from data the agency already holds. It ends with a severity-rated findings report, a backlog-ready issue log, consented highlight clips and a task metrics scorecard the program can repeat. The market band for a standalone round is \$15,000 to \$45,000 firm fixed price.

The illustration later in this paper works through the effort model. The base case is 8 participants over four weeks: 176 hours at a single \$110 to \$185 an hour market-norm rate, or \$19,360 to \$32,560, with a \$26K planning fee, well under the \$350,000 simplified acquisition threshold.

Avalon tests and advises. The study certifies nothing; it evidences. Compliance determinations, Paperwork Reduction Act decisions and acceptance of any product remain the agency's.

Why Services Fail Users Quietly

Any two of these phrases in one meeting usually mean the service has a usability problem nobody has measured.

What the program says	What it usually means
"Users keep calling the help desk about the same thing."	One step in the service fails often enough to generate its own call volume
"We get complaints but no data."	Feedback arrives as anecdote, so nobody can rank the problems
"We don't actually know our completion rate."	Analytics exist, but nobody has tied them to the tasks users came to do
"We launched and nobody's using it."	The service was built to the requirements document, not observed behavior
"Leadership wants proof before they'll fund the redesign."	The program needs evidence, not another opinion

Two routing questions keep this service distinct from its neighbors. If the complaint is about Section 508, WCAG or an accessibility conformance report, that is Avalon's separate Accessibility Modernization service, which tests against standards. If the question is about journeys across channels, CX governance or a measurement program, that is CX Strategy Development, also a separate Avalon service. This study watches real users attempt real tasks on one service.

Why Teams Skip the Evidence

- **The help desk already knows.** It knows what users report. Testing shows what users do, including the failures they never call about.
- **A survey seems enough.** A satisfaction score says something is wrong; it rarely says which step, or why.
- **More participants feel safer.** Five to eight participants in one segment typically surface the most frequent problems. Larger counts add cost and, with the public, can trigger the Paperwork Reduction Act.
- **The Paperwork Reduction Act seems to forbid it.** Lawful pathways exist, and the agency's PRA officer decides with a written plan in hand.
- **The incumbent offers to test its own work.** That is grading its own homework. An independent round gives the program a second set of eyes.

The Authorities and Where the Study Fits

Authority	What it says	What it means for this study
21st Century IDEA Act (2018)	Agencies modernize websites and digital services to be user-centered, accessible and consistent	The study supplies user research and feedback analysis the agency can cite in its digital experience reporting
Paperwork Reduction Act (44 U.S.C. ch. 35; 5 CFR 1320)	OMB clearance for collecting information from 10 or more members of the public; agencies often use generic clearances	A dependency: Avalon writes a PRA pathway memo, and the agency's PRA officer decides
Section 508 (2017 refresh incorporates WCAG 2.0 Level AA)	Federal ICT must be accessible; WCAG 2.1 and 2.2 are current W3C guidance	Adjacent only: usability findings are not conformance testing
U.S. Web Design System (USWDS)	Shared design components and guidance for federal websites	Findings reference USWDS patterns where the service uses them
FedRAMP Consolidated Rules (Jun 25, 2026)	Cloud offerings are "FedRAMP Certified" in classes A through D	Any SaaS research tool needs the agency's approval and a confirmed FedRAMP status; the base study needs none

One note on what this work is not. A usability study is not part of the Risk Management Framework, produces no authorization artifact, and does not contribute to an ATO. It sits beside security work, not inside it.

What the Study Does

The study covers one named service and one user segment, and works in the agency's own conferencing, file storage and backlog tools.

Workstream	Activities
Research planning	Research questions with the product owner, top-task scenarios, participant criteria, moderation guide, consent scripts and a PRA pathway memo
Participant recruitment	Six to eight participants from lists the agency supplies (program participants or employees), screened against the criteria and scheduled
Moderated sessions	One-hour remote sessions on the agency's approved platform, think-aloud protocol, with task success, errors, time on task and path recorded
Feedback analysis	Survey exports (such as Touchpoints), analytics funnels, site search terms and help desk ticket categories the agency already holds
Synthesis and prioritization	Affinity mapping, a severity rating for every finding, System Usability Scale (SUS) scoring where a benchmark helps, and a recommended fix with a rough level of effort
Readout and handoff	A findings briefing, consented highlight clips and an issue log formatted for the agency's tracker (Jira, GitHub or Azure DevOps)

Tasks, Participants and Sessions

The statement of work fixes four numbers: one segment, up to 8 participants, up to 5 tasks and 60-minute sessions. Tasks come from what users actually come to do, confirmed against the agency's analytics and help desk data, not from a feature list. Staff may stand in for the public only when the report labels them as proxies. Each finding cites the session and timestamp where it was observed, so the program can check the evidence itself.

The Feedback Annex

Most agencies already collect more feedback than they read. The annex ranks ticket categories by volume, finds where analytics funnels drop, lists what users search for and fail to find, and summarizes survey comments by theme. It then cross-checks those

signals against what the sessions showed. Where the two agree, the program has a finding it can defend; where they disagree, the report says so.

Deliverables

Deliverable	Format
Research plan (questions, tasks, participant criteria, schedule)	Word or Markdown document
PRA pathway memo	Short memo for the agency's PRA officer
Moderation guide and consent scripts	Agency-reviewable templates
Findings report with a severity-rated issue inventory	Briefing deck and written report
Backlog-ready issue log	Spreadsheet, CSV or tracker import
Consented highlight clips	Short video files in agency storage
Task metrics scorecard (success, errors, time on task, SUS)	One page, repeatable across rounds
Feedback analysis annex	Written annex with charts

Every deliverable is produced to conform with Section 508, and every one belongs to the agency.

Where the Study Stops

The study stops at the backlog. Fixing what it finds is design or development work: Avalon's UI/UX Design service, whose usability and design-system audit is an expert review without participants, or the relevant development service. Section 508 conformance testing, accessibility conformance reports and remediation belong to Accessibility Modernization. Sessions with assistive technology users are not offered today. Journey mapping and CX governance belong to CX Strategy Development, and navigation and labeling studies to Information Architecture, both separate Avalon

services. Survey research meant to be statistically generalizable, load, performance or security testing, formal user acceptance testing sign-off and license resale are out of scope.

What the Evidence Can and Cannot Claim

The accurate statement is narrow: this study provides user research, usability testing and customer feedback analysis of the kind the 21st Century IDEA Act directs agencies toward, with documentation suitable for the agency's digital experience and service delivery reporting. It does not make the service compliant with anything.

Accurate	Not accurate
Shows where real users fail real tasks, with evidence	Makes the service "IDEA-compliant"
Ranks problems by severity and rough fix effort	Makes the service "508-compliant"
Gives the program a repeatable task metrics baseline	Counts as Section 508 testing
An input to the product owner's priorities	Contributes to an ATO or accepts the product

Questions the Contracting Officer Will Ask

Question	Answer
Does the study need PRA clearance?	The agency's PRA officer decides. Avalon's memo lays out the plan, the participant count and the pathways; studies with 10 or more members of the public answering the same questions need OMB clearance, often under a generic clearance the agency already holds
Does it need IRB review?	Usability testing is usually operational, not research; the agency's own determination controls
Can participants be paid?	Only through a mechanism the agency's counsel confirms; the base study assumes no compensation
Who owns the data?	The agency owns all recordings, notes and findings; they stay in agency-approved storage and are transferred or destroyed as the SOW states
What tools are used?	Only tools the agency already authorizes; no new software is introduced

Consent language is reviewed by the agency's privacy office before the first session. Recordings are optional and can be limited to screen and audio. The findings report carries no personally identifiable information.

How a Study Runs

Phase	Timing	What happens
0: Kickoff and access	Award, in parallel	System access (staging preferred), participant list, privacy review of consent language; dependencies dated in writing
1: Plan and recruit	Week 1	Research questions, tasks, moderation guide, PRA pathway memo, screening and scheduling; feedback exports received
2: Sessions	Weeks 2–3	Moderated sessions, notes and coding; feedback annex analysis in parallel
3: Synthesis	Weeks 3–4	Affinity mapping, severity ratings, SUS, task metrics scorecard
4: Readout and handoff	Week 4	Briefing, clips, issue log imported into the agency's tracker

A second user segment adds one to two weeks and is priced as a change order. The scaled form of this service is a cadence, a round per release, not a bigger project. The delivery clock starts when access and participants are ready, because recruitment and access are the two dependencies that slip timelines.

What the Agency Provides

- A product owner authorized to approve tasks and receive findings, for two to four hours a week.
- Access to the service under test, staging preferred, including any PIV or account provisioning, dated at award.
- A participant list, or approval of how participants are invited, and a dated recruiting window.
- Privacy or legal review of consent language before the first session.
- The PRA officer's determination where members of the public take part.
- Exports of the survey, analytics, site search and help desk data the annex uses.

What Is Not Included

Included	Separate work
One segment, up to eight participants	Added segments, priced as change orders
Severity-rated findings and fix recommendations	Designing or building the fixes
Usability observations	Section 508 conformance testing or assistive technology sessions
Facilitating sessions	User acceptance testing sign-off
Unclassified services	Classified or cleared environments

How Avalon Staffs It

Avalon takes one engagement at a time, works business hours, and works on unclassified services only. Avalon's founder leads the engagement, manages the project and takes notes. A named UX researcher, subcontracted for the engagement, moderates every session; that researcher is identified and rate-locked before the proposal and listed in it with a résumé. Avalon has no past federal engagements under this service. FAR 15.305 says an offeror without a record of past performance may not be evaluated favorably or unfavorably on that factor; what protects the program is a named researcher, a fixed price, a small scope and deliverables the agency can inspect.

Requirement	How it is handled
UX researcher	Named personnel whose experience and investigation level match the agency's requirements, listed in the proposal
Access to internal systems	Public Trust investigation at the tier the agency requires, as a condition of the engagement
Tools	The agency's conferencing, storage and tracker, with access the agency provisions; no research platform licenses are needed or resold
Participants	Supplied or approved by the agency; Avalon holds no research panel
Insurance and contract vehicles	Any coverage the solicitation requires is a condition of award; Avalon does not hold a GSA Schedule or a governmentwide contract today

Key personnel substitution requires CO approval. The researcher never evaluates work Avalon designed or built.

Pricing and How to Buy

Engagement	Market-norm range	Avalon
Standalone moderated round: one service, one segment, 6–8 participants, about 4 weeks	\$15K–\$45K	Firm fixed price; illustrative \$26K
Multi-segment round	\$40K–\$85K	Added segments by change order, not the entry offer
Research inside a design or modernization task order	8–15% of that order's level of effort	Priced at the same researcher rate

The study sits under the \$350,000 simplified acquisition threshold and can be bought as an open-market purchase. Research inside a larger digital services order can be

subcontracted to Avalon by a prime. Avalon does not hold a GSA Schedule or a governmentwide contract today, and it does not resell software, licenses or recruiting services.

Usability work on a fielded service is usually funded from operations and maintenance money, which makes a four-week, fixed-price study a natural August or September purchase. Prices are fixed against one named service, a capped participant, task and session count, and a dated dependency schedule.

Illustration: One Service, Eight Participants, Four Weeks

The figures below are a planning illustration for one representative unclassified service: a public-facing online application for a federal benefit program, tested with first-time applicants, recruited from program participants the agency supplies. They are not a quote and not a result from any engagement.

Labor Loading

Planning, the moderation guide, consent scripts and the PRA memo take about 24 researcher and 8 analyst hours. Recruitment coordination takes 8 analyst hours plus one per participant. Each participant takes 1.5 hours of moderation and 1.5 of note-taking, then 1.5 hours each of coding. The feedback annex takes 8 researcher and 16 analyst hours, synthesis 16 and 8, and the readout and tracker import 8 and 8. Project management is 16 hours, about one-tenth time. Every role is priced in the same \$110 to \$185 band.

Role	Rate band	Base hours
Lead UX researcher	\$110–\$185/hr	80
Research analyst and note-taker	\$110–\$185/hr	80
Project manager	\$110–\$185/hr	16
Total		176

What the Study Costs

Scope	Hours	Rate-band range	Planning fee	Inside \$15K–\$45K
Lean (6 participants)	162	\$17.8K–\$30.0K	\$24K	Yes
Base (8 participants)	176	\$19.4K–\$32.6K	\$26K	Yes
Base plus contractor-assisted scheduling	192	\$21.1K–\$35.5K	\$29K	Yes

The base case is 176 hours, or \$19,360 to \$32,560, with a \$26K planning fee at the midpoint. The lean case with six participants runs \$17,820 to \$29,970. If the agency wants Avalon to handle scheduling and reminders, add 16 analyst hours: \$21,120 to \$35,520. Participant compensation and recruiting-vendor fees are excluded. A slip in access or recruiting does not change the price; it moves the clock.

As a cross-check, loading the round at 0.5 to 0.75 of a researcher, 0.25 to 0.5 of an analyst and 0.1 of a project manager for four weeks gives 136 to 216 hours, or \$14,960 to \$39,960. The base case sits in the middle. A five-week round at full loading and the top rate would reach \$49,950, above the \$45,000 ceiling, which is why the scope holds to four weeks. Each participant added within the segment costs about \$770 to \$1,295.

What the Agency Brings to the Readout

The model prices the study; it claims no saving. The size of the problem comes from the agency's own numbers: calls a month about the failing step, times the minutes each call takes, times twelve. At an illustrative 1,500 calls a month and eight minutes each, that is 2,400 staff hours a year, about 1.4 full-time staff. Add applications abandoned and paper fallbacks processed. Those figures, not a vendor estimate, go to the budget office.

Common Pitfalls

Pitfall	How to avoid it
"While you're at it, test the other audience too"	One segment, fixed participant, task and session counts; added segments by change order
Treating the study as a focus group	Participants attempt tasks; opinions are noted but findings rest on what they did
Staff standing in for the public without a label	Proxies are allowed only when the report says so
Recruiting late	Participant list and recruiting window dated in the SOW; the clock starts at readiness
Weak consent	Agency privacy review of the consent script before the first session
Compliance oversell	No claim that the study makes a service compliant, counts as Section 508 testing or supports an ATO

Conclusion and Next Step

Most struggling digital services are not missing a redesign. They are missing a record of what users actually do, ranked by how much it matters.

The natural first step is the fixed-price usability study: one service, one segment, about four weeks, under the simplified acquisition threshold, ending in a severity-rated findings report, a backlog the program can import, consented clips and a scorecard it can repeat after every release. To discuss a service, contact Avalon Technology Services at contact@AvalonTechServices.com.

Appendices

Appendix A: Glossary

Moderated usability testing

Sessions in which a researcher watches a participant attempt defined tasks and asks them to think aloud.

PRA pathway memo

A short memo describing the study's participants and questions so the agency's PRA officer can decide whether and how the Paperwork Reduction Act applies.

Severity rating

A defined scale ranking each finding by how badly it blocks or slows users.

SUS

System Usability Scale: a ten-item questionnaire that gives a repeatable usability score.

Top task

One of the few tasks most users come to a service to complete.

Appendix B: Illustration Assumptions

All values are planning assumptions for one representative unclassified service. None is drawn from a specific agency, a real service or a completed engagement.

Assumption	Value	Basis
Service	One public-facing online benefit application	Planning assumption
Scope caps	1 segment, 8 participants, 5 tasks, 60-minute sessions	Planning assumption
Per-participant effort	7 hours across recruiting, moderation, notes and coding	Planning assumption
Fixed effort	Researcher 56, analyst 48, PM 16 hours	KB delivery model; hours assumed
Rate	\$110–\$185 an hour for every role; \$150 planning midpoint	Market-norm labor band
Exclusions	Participant compensation, recruiting-vendor fees, travel	Not priced
Price check	Base fee \$26K	Inside \$15K–\$45K; under the SAT

Appendix C: References

1. **U.S. Congress**, *21st Century Integrated Digital Experience Act (H.R. 5759, 115th Congress)*
<https://www.congress.gov/bill/115th-congress/house-bill/5759>
2. **Office of the Law Revision Counsel**, *44 U.S.C. Chapter 35, Coordination of Federal Information Policy (Paperwork Reduction Act)*
<https://uscode.house.gov/view.xhtml?path=/prelim@title44/chapter35&edition=prelim>
3. **Electronic Code of Federal Regulations**, *5 CFR Part 1320, Controlling Paperwork Burdens on the Public*
<https://www.ecfr.gov/current/title-5/chapter-III/subchapter-B/part-1320>
4. **GSA Digital.gov**, *Paperwork Reduction Act guide*
<https://pra.digital.gov/>
5. **GSA Digital.gov**, *Touchpoints*
<https://touchpoints.digital.gov/>
6. **GSA Digital.gov**, *Digital Analytics Program*
<https://digital.gov/guides/dap/>
7. **GSA**, *U.S. Web Design System*
<https://designsystem.digital.gov/>
8. **GSA**, *Section508.gov*
<https://www.section508.gov/>
9. **U.S. Access Board**, *Information and Communication Technology (Revised 508 Standards)*
<https://www.access-board.gov/ict/>
10. **World Wide Web Consortium**, *Web Content Accessibility Guidelines (WCAG) 2.2*
<https://www.w3.org/TR/WCAG22/>
11. **FedRAMP**, *FedRAMP program and 2026 Consolidated Rules*
<https://www.fedramp.gov>
12. **Acquisition.gov**, *FAR 15.305, Proposal evaluation*
<https://www.acquisition.gov/far/15.305>

13. **Centre Law & Consulting**, *Inflation Updates to FAR Thresholds: What Changed on October 1, 2025*
<https://centrelawgroup.com/blog/inflation-updates-to-far-thresholds-what-changed-on-october-1-2025/>