Renaissance Philanthropy

Rapport: Internationale benchmark van gecoördineerde onderzoeksprogramma's

Rapport ingezonden door Renaissance Philanthropy voor Ministerie van Economische Zaken, Rijksoverheid
Januari 2026

Samenvatting voor het management

Dit rapport biedt een gevalideerde vergelijkende benchmark van zes toonaangevende agentschappen, geïnspireerd op het Advanced Research Projects Agency (ARPA)-model – DARPA, ARPA-E, ARPA-H, IARPA, ARIA en SPRIND – ter ondersteuning van de verkenning van een Nationaal Agentschap voor Disruptieve Innovatie (NADI) in Nederland. De benchmark richt zich op de werkingsmechanismen die verklaren hoe deze organisaties in de praktijk functioneren: hoe ze hun "recht om mee te doen" definiëren, programmadirecteuren/managers (PD's/PM's) bevoegdheden geven, instrumenten selecteren en combineren, en de transitie en implementatie vanaf het begin vormgeven. Het doel is niet om één enkel model voor te schrijven, maar om duidelijkheid te scheppen in de ontwerpkeuzes en afwegingen van belang.

Wat ARPA-achtige agentschappen gemeen hebben

1. Een scherp omschreven missie met expliciete randvoorwaarden ("right-to-play"-filters) die voorkomen dat het onderzoek afglijdt naar incrementeel of diffuus onderzoek en ontwikkeling.
2. Een programmatisch werkmodel gebouwd rondom geactiveerde PD's/PM's die programma's van begin tot eind ontwerpen en actief beheren tijdens de uitvoering – waarbij ze programma's beëindigen, bijsturen en intensiveren naarmate er nieuw bewijs naar voren komt.
3. Flexibele instrumentkeuze afgestemd op de programmalogica (in plaats van één dominant instrument).
4. De transitie en adoptie worden vanaf het begin zorgvuldig gepland (en niet tot het einde uitgesteld), met expliciete aandacht voor wie de resulterende functionaliteit moet adopteren, kopen, reguleren of beheren om ermee impact te genereren.

Belangrijkste bevindingen

1. Het "ARPA-model" is in de eerste plaats een operationele methodologie, geen sector. De sector waarop het wordt toegepast (defensie, energie, gezondheidszorg, inlichtingendiensten, de gehele economie) bepaalt het transitietraject meer dan dat het de kern van de managementlogica verandert.
2. Autonomie voor PD's/PM's is een minimale vereiste: wanneer beslissingsbevoegdheden teruggegeven worden aan commissies, vallen agentschappen terug op conventionele subsidieverlening en verliezen ze de mogelijkheid om tijdsgebonden, op mijlpalen gebaseerde ondernemingen uit te voeren.
3. Succesvolle modellen maken onderscheid tussen (i) de toestemming om een programma te starten en (ii) de bevoegdheid om het te beheren zodra het is gelanceerd. De meeste modellen bieden een lichtgewicht goedkeuringsprocedure voor de lancering, gevolgd door een hoge mate van autonomie binnen de goedgekeurde kaders.
4. Bij 'kill/scale' gaat het meestal om herverdeling binnen een portfolio: zwakkere benaderingen worden geschrapt zodat middelen kunnen worden ingezet voor sterkere benaderingen, in plaats

van dat ze 'verloren' gaan binnen het programma. Een geloofwaardige 'kill/scale' vereist meetbare mijlpalen en een regelmatige besluitvorming.

5. De transitie kan niet worden uitgesteld tot het einde van een programma. De sterkste transitiesystemen betrekken gebruikers/kopers en testen een "plausibel traject" al vroeg in het proces: ARPA-E's Tech-to-Market (T2M)-beoordeling; DARPA's logica voor transitiepartners; IARPA's vroege klantbetrokkenheid; ARPA-H's 'systeemontwikkeling' voor regelgeving en vergoedingen.
6. De keuze van het instrument bepaalt wie kan deelnemen en hoe snel het agentschap kan innoveren. Waar wettelijke kaders de instrumenten beperken (met name op het gebied van aandelenkapitaal en aanbesteding), ontwikkelen agentschappen ofwel conforme trajecten (bijvoorbeeld de co-investerings- en licentiestrategie van SPRIND) of compenseren ze via andere middelen (bijvoorbeeld de flexibiliteit van ARIA bij de aanbesteding; de architectuur van ARPA-E met ondersteuning en commercialisering).
7. De Europese institutionele context is van belang. Met name SPRIND en ARIA bieden relevante lessen voor NADI over autonomie, aanbestedingen en beperkingen op staatssteun, en over hoe de flexibiliteit in een parlementaire omgeving te behouden.
8. Een gering personeelsbestand met sterke ondersteunende functies is een consistent patroon. In alle modellen wordt een kleine kern van PD's/PM's omringd door deskundige ondersteunende functies (contractering/juridische zaken, financiën, testen en evaluatie (T&E)/validatie, transitie/commercialisering), vaak aangevuld met externe partijen of partnerorganisaties.
9. De grootste variatie binnen de verschillende domeinen zit niet in de vraag of programma's worden geleid door PD/PM-teams en dat er mijlpalen worden bereikt (dat patroon blijft hetzelfde), maar in de belangrijkste toegangspoort tot implementatie (aanbestedingsinstantie, toezichthouder/betaler of de markt) en daarmee in het soort bewijsmateriaal van transitieniveau dat programma's moeten produceren om daadwerkelijke implementatie te bewerkstelligen.
10. De lanceringsfase is een ontwerpprobleem: ARIA en SPRIND laten zien dat vroege beperkingen op het gebied van werving/contractering padafhankelijkheid kunnen creëren sturend op langzamere, op commissies gebaseerde werkwijzen; omgekeerd kan een klein, bevoegd oprichtingsteam met duidelijke taakverdelingen legitimiteit opbouwen en tegelijkertijd snelheid behouden.

Implicaties voor NADI: ontwerpkeuzes om expliciet te maken

1. Hoe streng moeten de "ARPA-waardige" filters van NADI zijn (focus op de missie versus breedte)?
2. Welke rechtsvorm en bevoegdheden zijn vereist zodat NADI (a) termijngebonden PD's/PM's concurrerend kan aannemen, (b) snel contracten kan afsluiten met mijlpaalcontrole, en (c) beslissingsrechten kan beschermen tegen comitébestuur?
3. Voor elk domein waarin NADI actief is, wat moet de standaard overgangslogica zijn: overheid als koper, markt als koper, of een combinatie hiervan – en welke minimale betrokkenheid van de koper (of ecosysteemcapaciteit) is vereist vóór de lancering?
4. Welke instrumenten zijn haalbaar en wenselijk onder Nederlands/EU-recht (subsidies, contracten, pre-commerciële aanbesteding/innovatieve openbare aanbesteding (PCP/PPI), prijzen, aandelen/quasi-aandelen), en welke positie op het gebied van intellectuele eigendomsrechten/gegevensrechten ondersteunt de transitie met behoud van het algemeen belang?
    a. Zijn er nieuwe wettelijke machtigingen of uitzonderingen nodig om dit te bereiken?

5. Welke ondersteunende functies moeten vanaf dag één aanwezig zijn (met name op het gebied van contracten/juridische zaken en transitie) om te voorkomen dat er achteraf capaciteitstekorten ontstaan?
6. Hoe zal NADI de paradox van de opstartvolgorde oplossen: wat moet er vóór de lancering door ambtenaren worden opgezet (juridische/HR/inkoopstructuur) en wat moet bewust worden overgelaten aan de toekomstige NADI-directeur en de eerste programmadirecteuren/projectmanagers (portfoliokeuzes, aanwervingsprioriteiten, programmaselectie), om in een vroeg stadium vast te zitten aan processen te voorkomen?

Overzicht van agentschappen (archetypen op hoog niveau)

Bureau	Missieanker	'Poolster' transitie	Sterke punten	Spanning/risico
DARPA	Technologische verrassing defensie	Transitie naar militaire gebruikers / inkoopprogramma's	Sterke autonomie van de projectmanager; brede flexibiliteit van instrumenten; duidelijke gebruikersankerpunten.	De implementatie in latere fasen is afhankelijk van de sponsors van de dienst en de inkoopprocessen buiten DARPA.
ARPA-E	Doorbraken in energietechnologie	Commerciële uitrol via de industrie/markten; incidenteel vervolg in de publieke sector	T2M ingebouwd; mijlpaalbeheer; positionering in 'lege ruimte'	Minder directe invloed van inkoop/de koper; de uitkomst is afhankelijk van het kapitaal en de regelgeving.
ARPA-H	Transformatieve gezondheidsresultaten	Adoptie via gezondheidszorgsystemen, regelgevende instanties, zorgverzekeraars en/of commercialisering.	Breed scala aan wettelijke instrumenten; flexibiliteit gericht op andere transacties (OT); expliciete engineering van adoptie.	Complexe transitiebarrières (regelgeving, vergoedingen, werkprocessen) kunnen de technologische vooruitgang belemmeren.
IARPA	Intelligentievoordeel	Overgang naar 'klanten' binnen de inlichtingengemeenschap	Geïnstitutionaliseerde onafhankelijke T&E; strenge meetmethoden; probleemformulering over instantiebarrières	Geen interne implementatie-eenheid; de overgang is afhankelijk van externe inlichtingenpartners die de resultaten oppakken.
ARIA	Wetenschap en uitvindingen met een hoog risico en een hoge opbrengst.	Integratie door het ecosysteem; ondernemerschapstrajecten	Grote autonomie onder de ARIA-wet; flexibiliteit bij de opdrachtverlening; tolerantie voor vernieuwing.	Risico op fragmentatie zonder sterke grensregels; transitiemechanismen nog in ontwikkeling.
SPRIND	Sprong voorwaarts in innovaties; strategische/soevereiniteitskadering	Ondernemingsontwikkeling + marktgroei; enkele trajecten binnen de publieke sector	Draaiboek voor de Europese rechtscontext; gefaseerde uitdagingen; de Freedom Act heeft de autonomie en instrumenten uitgebreid.	Beperkingen op het gebied van staatssteun/aanbesteding en vereisen een zorgvuldige structurering; vangrails op co-investeringen en licenties zijn nodig.

Renaissance Philanthropy

Report: International Benchmark of Coordinated Research Programmes

Report produced by Renaissance Philanthropy for Ministry of Economic Affairs, Government of the Netherlands
January 2026

Executive summary

This report provides a validated comparative benchmark of six leading agencies inspired by the Advanced Research Projects Agency (ARPA) model — DARPA, ARPA-E, ARPA-H, IARPA, ARIA, and SPRIND — to support the exploration of a National Agency for Disruptive Innovation (NADI) in the Netherlands.

The benchmark focuses on the operating mechanisms that explain how these organisations function in practice: how they define their "right-to-play," empower Programme Directors/Managers (PDs/PMs), select and combine instruments, and engineer transition and adoption from the outset. The aim is not to prescribe a single model, but to clarify design choices and trade-offs that matter.

What ARPA-type agencies have in common

1. A sharply defined mission with explicit boundary rules ("right-to-play" filters) that prevent drift into incremental or diffuse R&D.
2. A programmatic operating model built around empowered Programme Directors (PDs) / Programme Managers (PMs) who design programmes end-to-end and actively manage them throughout execution — terminating, redirecting, and doubling down as evidence emerges.
3. Flexible instrument choice matched to programme logic, rather than a single dominant tool.
4. Transition and adoption are engineered from the outset (not left for the end), with explicit attention to who needs to adopt, buy, regulate, or operate the resulting capability in order to ensure that it translates into impact.

Key findings

1. The "ARPA model" is primarily an operating methodology, not a sector. The sector to which it's applied (defence, energy, health, intelligence, economy-wide) shapes the transition pathway more than it changes the core management logic.
2. PD/PM autonomy is a lower-bound requirement. When decision rights are pulled back into committees agencies revert to conventional grant-making and constrain their ability to run time-bound, milestone-driven bets.
3. Successful models distinguish between (i) authorisation to start a programme and (ii) authority to manage it once launched. Most have lightweight approval to launch, followed by high autonomy within the approved envelope.
4. Kill/scale is usually about within-portfolio reallocation: weaker approaches are cut so resources can be shifted to stronger approaches within the same programme, rather than being 'lost' entirely. Credible kill/scale requires measurable milestones and a clear decision-making cadence.
5. Transition cannot be postponed to the end of a programme. The strongest transition systems put adopter/buyer engagement and "plausible pathway" tests upstream — ARPA-E's Tech-to-Market

(T2M) gating; DARPA's transition partner logic; IARPA's early customer involvement; ARPA-H's regulatory/reimbursement 'system engineering'.

6. Instrument choice determines who can participate and how quickly the agency can iterate. Where legal frameworks constrain instruments (especially equity and procurement), agencies either build compliant pathways (e.g., SPRIND's co-investment and licensing posture) or compensate via other levers (e.g., ARIA's commissioning flexibility; ARPA-E's assistance + commercialisation architecture).
7. European institutional context matters. SPRIND and ARIA offer especially relevant lessons for NADI on autonomy, procurement and state-aid constraints, and how to preserve agility in a parliamentary setting.
8. Lean headcount with strong enabling support is a consistent pattern. Across models, a small core of PDs/PMs is wrapped by expert enabling functions (e.g. contracting, legal, finance, Test & Evaluation (T&E), commercialisation), often supplemented by contractors or partner organisations.
9. Across domains, the biggest variation is not in whether programmes are PD/PM-led and milestone-gated (that pattern holds), but in the dominant adoption gate (procurement authority, regulator/payer, or market) and therefore the kind of transition-grade evidence programmes must produce to trigger real uptake.
10. Launch phases are a design problem: ARIA and SPRIND show that a small, empowered founding team with clear delegations can build legitimacy while preserving speed; conversely, early constraints on hiring/contracting can create path-dependence toward slower, committee-oriented practices.

Implications for NADI: design choices to make explicit

1. How tight should NADI's "ARPA-worthy" filters be (mission focus vs breadth)?
2. What legal form and delegations are required so NADI can (a) hire term-limited PDs/PMs competitively, (b) contract rapidly with milestone control, and (c) protect decision rights from committee governance?
3. For each domain that NADI operates in, what should the default transition logic be — government-as-buyer, market-as-buyer, or mixed — and what minimum buyer commitment or ecosystem capability is needed before launch?
4. What instrument toolkit is feasible and desirable under Dutch/EU law (grants, contracts, Pre-Commercial Procurement / Public Procurement of Innovative solutions (PCP/PPI), prizes, equity/quasi-equity), and what IP/data rights posture will support transition while maintaining public benefit?
    a. Are any new legal capabilities or exemptions required to achieve this?
5. What enabling functions must exist on day one (especially contracting/legal and transition) to prevent 'afterthought' capability gaps?
6. How will NADI resolve the start-up sequencing paradox: what must be set up by civil servants pre-launch (legal/HR/procurement scaffolding) versus what should be deliberately left to the future NADI Director and early PDs/PMs (portfolio choices, hiring priorities, programme selection) to avoid unintentional early process lock-in?

Renaissance Philanthropy

Agencies at glance (high-level archetypes)

Agency	Mission anchor	Transition 'north star'	Strengths	Tension/risk
DARPA	Defence technological surprise	Transition to military users / acquisition programmes	Strong PM autonomy; broad instrument flexibility; clear user anchor	Downstream adoption depends on service sponsors and acquisition pathways outside DARPA
ARPA-E	Energy technology breakthroughs	Commercial deployment via industry/markets; occasional public follow-on	T2M built in; milestone management; 'white space' positioning	Less direct procurement/buyer leverage; outcomes depend on capital/regulatory environment
ARPA-H	Transformative health outcomes	Adoption via health systems, regulators, payers, and/or commercialisation	Broad statutory tool menu; Other Transaction (OT) centric flexibility; explicit engineering of adoption	Complex transition barriers (regulatory, reimbursement, workflows) can dominate technical progress
IARPA	Intelligence advantage	Transition to Intelligence Community (IC) 'customers'	Institutionalised independent T&E; rigorous metrics; cross-agency problem framing	No internal deployment arm; transition depends on external IC partners picking up results
ARIA	High-risk, high-reward science & invention	Ecosystem uptake; entrepreneurship pathways	High autonomy under ARIA Act; commissioning flexibility; tolerance for novelty	Risk of fragmentation without strong boundary rules; transition mechanisms still maturing
SPRIND	Leap innovations; strategic / sovereignty framing	Venture creation + market scale; some public-sector pathways	European-legal-context playbook; staged challenges; Freedom Act expanded autonomy and instruments	State-aid/procurement constraints require careful structuring; needs co-investment and licensing guardrails

Renaissance Philanthropy

1. Purpose, scope, and method

1.1 Purpose

The Government of the Netherlands is exploring the creation of an ARPA-like capability to accelerate "disruptive innovation" through coordinated, time-bound research and development programmes. This benchmark supports that design effort by comparing six internationally prominent ARPA-type agencies across a common set of dimensions, with a focus on operational mechanisms (decision rights, cadence, instruments, and transition pathways).

1.2 Scope of benchmark

- Agencies: DARPA, ARPA-E, ARPA-H, IARPA, ARIA, SPRIND.
- Unit of analysis: the agency operating model (not individual programme performance evaluation).
- Primary lens: what design features are necessary to run coordinated, high-risk, time-bound programmes, what varies, and why.

1.3 Methods and evidence base

We used structured desk research and document analysis, organised around a comparative coding framework. For each agency and dimension, we synthesised primary sources (statutes, official guidance, annual reports, budget documents, procurement/contracting guidance) and secondary sources (evaluations, academic analyses, credible practitioner commentary).

This approach involved:

- Structured comparison: a common matrix (dimensions A–I) populated for each agency.
- Triangulation: where feasible, major claims were cross-checked using multiple independent sources, including interviews with insiders, key opinion leaders, and former ARPA employees.
- Limitations: budgets, staffing, and programme counts are not always reported in comparable formats; some agencies publish more detail than others; classified or sensitive programmes (notably in IARPA) limit public comparability. Confidence: where the evidence base is thinner, we label findings as indicative and provide the most relevant public references for follow-up.

1.4 How to read this report

Sections 2–3 provide snapshots of agencies and comparative findings across the requested dimensions. Section 4 distils cross-cutting design choices for NADI and presents decision questions rather than prescriptive recommendations. Finally, section 5 sketches a possible staged implementation for NADI and section 6 provides concluding remarks.

A separate Evidence Pack accompanies this report and includes the underlying matrices, glossary, methods note, and reference library.

2. Agency snapshots

The agencies benchmarked share a recognisable ARPA-style operating logic but differ in their mission anchors, legal forms, instruments, and transition environments. These snapshots highlight the aspects that most strongly condition the rest of the operating model.

DARPA (United States, Department of Defence)

Mission: prevent and create technological surprise for national security; time-bound, breakthrough R&D programmes executed via external performers.

- **Organisational logic:** empowered PMs propose and run programmes with a flat hierarchy (PM → Office Director → Director).
- **Typical programme:** 3–5 years, multiple performers, aggressive milestones. Transition to a defence customer is actively engineered but typically executed outside DARPA.
- **Instruments:** Broad Agency Announcements (BAAs) leading to contracts, grants/cooperative agreements, and Other Transaction (OT) instruments for prototypes.
- **Scale:** small core workforce (often cited at ~200–220 government employees) supporting ~100 PMs and a large external performer base.

Illustrative recent programmes:

- **AI Cyber Challenge (AIxCC, 2023–2025):** DARPA ran a multi-stage, prize-driven competition (with a public finale at DEF CON 33) to accelerate AI systems capable of finding and patching vulnerabilities in widely used open-source software.
    - *What it illustrates*: how DARPA uses prizes and staged down-selects to crowd in non-traditional teams, create a measurable evaluation arena, and build momentum in an emerging capability area. (https://www.darpa.mil/research/programmes/ai-cyber)

- **Air Combat Evolution (ACE, 2019–2024):** Progressed from simulation to flight testing on a modified F-16 (X-62A VISTA), generating transition-relevant evidence about autonomy and human-machine teaming.
    - *What it illustrates*: DARPA's preference for operationally grounded demonstrations with partners to de-risk adoption decisions, even when DARPA is not the downstream buyer. (https://www.darpa.mil/research/programmes/air-combat-evolution)

- **Blackjack (2017–present):** Develops and demonstrates a resilient low-Earth-orbit satellite network by integrating commercial smallsat advances into defence-relevant architectures.
    - *What it illustrates*: DARPA's role in prototyping and de-risking architectures that can later move into service acquisition pathways, with early attention to interoperability and operations concepts. (https://www.darpa.mil/research/programmes/blackjack)

ARPA-E (United States, Department of Energy)

Mission: high-risk, high-reward energy technology innovations that are too early or risky for private investment; positioned as a gap-filler in the U.S. energy innovation ecosystem.

- **Distinctive feature:** a formal T2M function that embeds commercialisation planning and milestones into awards.

- **Programmes:** portfolio-managed; selection is not purely peer review — PDs curate a portfolio to balance risk and approach diversity.
- **Instruments:** primarily financial assistance (cooperative agreements) with milestone governance; selective bridges toward scale (e.g., SCALEUP).
- **Scale:** lean staff (~100) administering a programme budget on the order of hundreds of millions of dollars per year.

Illustrative recent programmes:

- **Vision OPEN 2024 (launched 2024):** An "open" funding opportunity inviting high-risk proposals aligned to three broad system-level goals (clean primary energy abundance; an intermodal energy 'superhighway'; and a carbon transition for materials).
    - *What it illustrates*: how ARPA-E uses a bounded mission frame to explore white-space opportunities without relying on a narrow topic list, while still structuring awards around technical and commercial milestones. (https://arpa-e.energy.gov/programmes-and-initiatives/view-all-programmes/vision-open-2024)

- **SCALEUP (commercialisation bridge; ongoing since 2019):** Follow-on awards to prior ARPA-E projects assessed to have a viable route to commercial deployment; designed to bridge the scale-up gap to first factories, pilots, or commercial products.
    - *What it illustrates*: ARPA-E's explicit T2M posture and use of award design (larger tickets; deployment-oriented milestones; commercial partnerships) as a policy lever. (https://arpa-e.energy.gov/programmes-and-initiatives/SCALEUP-programme)

- **Example performer pathway: Antora Energy (thermal batteries; SCALEUP award announced 2024):** Funding to accelerate pilot-scale production/manufacturing steps for a combined heat-and-power thermal battery product for industrial customers.
    - *What it illustrates*: the practical handoff from technology risk to manufacturing/deployment risk — and how ARPA-E can target that latter risk through milestone-based scale-up support. (https://arpa-e.energy.gov/news-and-events/news-and-insights/arpa-e-investor-update-vol-22-antora-energys-thermal-batteries)

ARPA-H (United States, Department of Health & Human Services)

Mission: accelerate health breakthroughs with outsized impact; operates across biomedical research, health systems innovation, and enabling platforms.

- **Distinctive feature:** broad statutory tool menu with strong emphasis on OTs and tailored terms (IP, data, milestones) suitable for complex health translation.
- **Transition logic:** adoption depends on health systems, payers, regulators, and clinical workflows; ARPA-H therefore emphasises 'system' transition planning.
- **Transparency regime:** subject to U.S. Freedom of Information Act (FOIA) with statutory protections for confidential commercial/financial information; annual reporting and external evaluation requirements.
- **Scale:** statute sets a headcount cap (210); early staffing has been lean relative to ambitions.

Illustrative recent programmes:

- **THEA – Transplantation of Human Eye Allografts (launched 2024):** Aims to enable whole functional eye transplantation and related regenerative technologies (e.g., preserving/regrowing nerves from eye to brain).
    - *What it illustrates*: an ARPA-style programme framed around an ambitious end-state with multiple technical thrusts, where "success" requires building a system of capabilities that can plausibly translate into clinical practice. (https://arpa-h.gov/explore-funding/programmes/thea)

- **DIGIHEALS (launched 2023):** Targets resilience of digital health infrastructure, aiming to prevent and mitigate cyberattacks on medical facilities and strengthen the electronic health ecosystem (including adapting proven technologies developed for national security).
    - *What it illustrates*: health transition constraints are often operational and institutional (workflows, IT systems, risk/compliance) rather than purely technical — requiring early engagement with adopters and "deployment-grade" evidence. (https://arpa-h.gov/explore-funding/programmes/digiheals)

- **PARADIGM (launched 2024):** Seeks a scalable mobile platform to deliver advanced medical services outside hospitals, particularly in rural settings.
    - *What it illustrates*: ARPA-H's emphasis on delivery-system innovation and on designing programmes around real-world adoption gates (providers, payers, regulators), not just prototyping a device or software artefact. (https://arpa-h.gov/explore-funding/programmes/paradigm)

IARPA (United States, Office of the Director of National Intelligence)

Mission: high-risk, high-payoff research for the U.S. Intelligence Community (IC), often focused on forward-looking capabilities.

- **Distinctive feature:** institutionalised independent T&E and rigorous metrics; in some programmes, a substantial budget share can be allocated to T&E.
- **Programmes:** commonly run at Technology Readiness Level (TRL) 3–5; phase-gated down-selects are common; transition partners are involved early, but adoption is not controlled.
- **Instruments:** BAAs typically lead to Federal Acquisition Regulation (FAR) based R&D contracts; the process often starts with white papers before full proposals.
- **Scale:** much smaller than DARPA—public sources often describe budgets in the 'few hundred million' range with ~20–30 programmes.

Illustrative recent programmes:

- **TrojAI (AI security; multi-year programme):** Develops techniques to detect and mitigate malicious "Trojan/backdoor" attacks on AI systems, with challenge rounds and a dedicated T&E function to compare approaches against defined metrics.
    - *What it illustrates*: IARPA's hallmark use of independent, metric-driven evaluation regimes ("bake-offs") to make progress legible to mission owners and to support objective down-selects. (https://www.iarpa.gov/research-programmes/trojai)

- **'Gold standard' independent T&E as an operating norm (cross-programme practice):** IARPA commonly reserves a substantial share of programme budget for an independent T&E team that validates performer results against programme metrics.
    - *What it illustrates*: in domains where deployment is sensitive (data rights, security constraints) and where reproducibility matters, evaluation capability is not an afterthought — it is part of the core programme design and budget. (https://www.iarpa.gov/images/pdfs/GSS.pdf)

ARIA (United Kingdom, Advanced Research and Invention Agency)

Mission: fund high-risk, high-reward research with autonomy to back unconventional approaches; created to complement existing UK R&D institutions.

- **Legal form:** non-departmental public body with high autonomy under the ARIA Act; ministers cannot direct individual funding decisions.
- **Distinctive feature:** broad freedom in commissioning and contracting, including the ability to use mechanisms not typical for standard grants.
- **Governance:** board + CEO; PDs are expected to actively manage portfolios and take risks.
- **Budget:** currently £220m/year rising to £400m/year by FY2029/2030

Illustrative recent programmes:

- **Sustained Viral Resilience (£46m, falling within the broader ARIA opportunity space: Sculpting Innate Immunity):** Pursues a new class of medicines that provide durable, broad-spectrum protection against respiratory viruses by engineering the innate immune system.
    - *What it illustrates*: ARIA's 'opportunity space' approach (a small number of coherent, high-risk bets) and its ability to run programmes that combine ambitious end goals with multiple technical avenues. (https://www.aria.org.uk/opportunity-spaces/sculpting-innate-immunity/sustained-viral-resilience/)

- **Precision Neurotechnologies (£69m, falling within the broader ARIA opportunity space: Scalable Neural Interfaces):** Targets new methods to interface with the human brain at the circuit level with unprecedented precision.
    - *What it illustrates*: ARIA's use of sizeable, time-bound programmes to pursue frontier capabilities, under a governance model intended to protect programme discretion while maintaining public legitimacy. (https://www.aria.org.uk/precision-neurotechnologies/)

SPRIND (Germany, Federal Agency for Disruptive Innovation)

Mission: identify and support 'sprunginnovationen' (leap innovations), including via challenge-based approaches and venture creation; positioned within EU legal constraints.

- **Legal form:** federally owned Gesellschaft mit beschränkter Haftung (GmbH) — a German limited-liability company form — with a supervisory board; the 2023 SPRIND Freedom Act expanded autonomy and flexibility (including hiring and investment freedoms).

- **Distinctive features:** staged challenges and sprints; ability to use equity/quasi-equity with co-investment requirements; ability to retain a share of returns for reinvestment.

- **State-aid posture:** IP typically remains with the inventors while the government retains a free, non-exclusive license; co-funding helps align with state-aid guidelines.
- **Budget:** ~€1bn over 10 years (indicative) with annual allocations varying by year.

Illustrative recent programmes:

- **Challenge: Broad-Spectrum Antivirals (winners announced 2024):** A staged challenge to identify breakthrough antiviral approaches; SPRIND continues to support promising projects even after the formal challenge ends.
    - *What it illustrates*: SPRIND's competition-driven, stage-gated model ("fund fast, down-select hard"), and its explicit focus on post-challenge continuation toward real-world deployment. (https://www.sprind.org/en/actions/challenges/antiviral)

- **Challenge: Carbon-to-Value (winners announced 2024):** Staged competition to develop methods to durably remove CO_2 and bind it in products; winners were selected by an expert jury, with continued support for breakthrough potential after the competition.
    - *What it illustrates*: a structured path from many approaches → a few validated demonstrations, coupled with SPRIND's venture-building posture to carry winners beyond the competition endpoint. (https://www.sprind.org/en/actions/challenges/carbon-to-value)

3. Benchmark findings

This section presents the comparative benchmark across the nine requested dimensions. For each dimension we describe: (i) cross-agency patterns that define the ARPA operating system; (ii) meaningful variants across the six agencies (often driven by domain and legal context); and (iii) decision questions relevant to NADI's design.

3.1 Mandates and governance

Cross-agency pattern

- **Autonomy with accountability:** ARPA-type agencies typically receive broad mission mandates and delegated authorities, paired with ex post oversight (budgets, audits, reporting, evaluations) rather than ex ante micro-approval of each award.
- **Independence safeguards:** mechanisms to protect technical decision-making from day-to-day political or bureaucratic interference are common, even when agencies are nested within large departments.
- **Lean governance layers:** decision rights commonly sit with PDs/PMs for programme execution, with only 1–2 management layers above them; committees are used for advice, not operational approvals.

Meaningful variants across agencies

- **Department-embedded vs arm's-length:** DARPA, ARPA-E, ARPA-H, and IARPA sit inside large U.S. departments/agencies, whereas ARIA (UK) and SPRIND (DE) were created with explicit legal separation and bespoke governance arrangements.
- **Transparency regimes:** ARIA has a statutory exemption from UK Freedom of Information (FOI), while U.S. agencies operate under FOIA with varying carve-outs (e.g., ARPA-H protections for confidential commercial/financial information).
- **Procurement constraints and exemptions:** ARIA is exempt from certain UK procurement rules; SPRIND operates under EU/German procurement and state-aid constraints but has special flexibilities under the Freedom Act.

Design questions for NADI

- What legal form best preserves operational autonomy while meeting Dutch accountability norms (ministerial responsibility, parliamentary scrutiny, audit requirements)?
- Which decisions must be protected as technical/programmatic judgment, and which can remain subject to administrative controls?
- What transparency posture is both politically viable and operationally workable (e.g., publication of awards, redaction of sensitive commercial information)?
- How will NADI demonstrate accountability for risk-taking (reporting, evaluation cadence, auditability) without introducing committee governance?

3.2 Challenge definition ("ARPA-worthy" problems)

Cross-agency pattern

- **Problem selection is a core competency**, not a one-time intake exercise. Agencies invest heavily in upstream problem framing and in testing whether a problem is (a) important, (b) not solvable through conventional R&D, and (c) tractable within a time-bound programme.

- **Many models use explicit rubrics** (e.g., Heilmeier-style questions) to clarify the objective, novelty, approach portfolio, and success metrics.
- **Boundary rules are essential:** agencies define what they do not fund (incremental work, routine scaling, undifferentiated basic research) to avoid dilution.

Meaningful variants across agencies

- **Customer pull vs market pull:** defence and intelligence ARPAs often have identifiable government 'customers' (services, agencies), while ARPA-E and parts of ARPA-H more often rely on market and ecosystem adoption (buyers, investors, regulators).
- **Top-down themes vs bottom-up PD/PM entrepreneurship:** some agencies emphasise strategic thrust areas (especially new agencies at launch), while mature models rely more heavily on PD/PM-initiated programmes within mission boundaries.
- **Time horizon and risk:** IARPA often accepts longer time-to-impact and earlier TRLs; ARPA-E explicitly screens for commercialisation plausibility; ARPA-H must integrate clinical and regulatory realities.

Design questions for NADI

- How will NADI define and operationalise 'disruptive innovation' in a way that is selective enough to protect scarce PD/PM attention?
- What is the right balance between national priorities and PD/PM-initiated programme ideas?
- What minimum evidence should be required before launch (e.g., a transition hypothesis; credible adopter/buyer map; regulatory pathway where relevant)?

3.3 Programme anatomy (scope, TRL range, ticket size, kill/scale, PD/PM load)

Cross-agency pattern

- **Programmes are time-bound and milestone-driven** (often 3–5 years as a dominant pattern), with explicit go/no-go gates and down-selects.
- **Programmes are portfolios:** multiple technical approaches are funded in parallel early, then narrowed as evidence accumulates.
- **PDs/PMs actively manage programmes throughout (not just at launch):** frequent reviews, site visits, renegotiation of milestones, and reallocation of funds are routine.

Indicative quantitative parameters (where available)

- DARPA: PM tenure is typically under ~5 years; PMs often run several programmes concurrently, each being 3–5 years; example PM portfolios can be tens of millions of dollars per year.
- IARPA: work is often at TRL 3–5; PMs commonly manage ~2 programmes concurrently; typical programme scale can be roughly $20–50M over 3–5 years (public, indicative).
- SPRIND (example challenge): staged funding can start at around €0.7M per team (year 1), scaling to ~€1.5M (year 2) and ~€2.5M (year 3) for teams that advance.

Meaningful variants across agencies

- **Independent test and evaluation:** IARPA institutionalises independent T&E as a distinct function with budget share, strengthening kill/scale credibility and scientific truth.

- **Demonstration and prototyping:** DARPA often drives toward prototype demonstrations relevant to defence users; ARPA-E varies by technology but often aims for lab-to-pilot readiness; health programmes may require validation, clinical pathways, and data infrastructure.
- **PD/PM bandwidth:** where programmes involve sensitive data, complex contracting, or multi-stakeholder adoption (health), PD/PM load may be lower unless there's enabling support.

Design questions for NADI

- What is NADI's target TRL 'sweet spot' by programme type (e.g., proof-of-concept vs prototype vs pre-commercial scale), and how does that vary by theme?
- What are the standard artefacts for programme governance (charter, milestones/metrics, transition plan, independent validation plan)?
- What kill/scale authority will PDs/PMs have in practice (budget reallocation, termination, downselects), and what is the cadence of decision-making?
- Given Dutch capacity and legal constraints, what is a realistic PD/PM workload (number of active programmes and project performers per PD/PM) and what support is required?

Across the benchmark, the most transferable 'programme anatomy' is a repeatable lifecycle with explicit decision gates. Below is a neutral reference model you can adapt by domain.

Stage	Primary objective	Key decisions	Typical artifacts	Failure mode
Opportunity framing (weeks)	Define the problem, the right-to-play fit, and the success criteria.	PD/PM drafts; Director/Office approves to proceed (lightweight gate).	One-page concept; mission filter; adopter map; draft metrics; risks screen.	Problem too broad; metrics not falsifiable; no credible transition hypothesis.
Programme design (4-10 weeks)	Design portfolio approach and learning plan (how uncertainty will be resolved).	PD/PM owns design; internal review tests clarity, tractability, and transition realism.	Programme charter; milestones; evaluation plan; contracting approach; IP posture.	The programme becomes a list of topics; adoption constraints are not integrated early.
Performer sourcing and selection (4-12 weeks)	Find diverse approaches and assemble teams (including non-traditional performers).	PD/PM owns technical selection within remit; contracting/legal enables fast, compliant awards.	Solicitation package; short concept intake; selection rationale; negotiation checklist; COI log.	Slow cycle time; over-engineered panels; inability to contract with startups/SMEs.
Phase 1: exploration (6-18 months)	Run multiple approaches in parallel; generate comparable evidence; downselect.	PD/PM owns keep/redirect/stop at set cadence; leadership reviews portfolio health.	Quarterly reviews; milestone reports; updated transition plan; test results; updated risk notes.	Milestones too soft to kill; evidence not comparable; reallocation does not happen.
Phase 2: demonstration (6-18 months)	Prove system-level feasibility in relevant conditions; package evidence for decision-makers.	PD/PM owns integration; transition partners participate at hinge points without taking control.	Demo plan; testbed access; validation package; compliance artefacts; cost/impact model.	Prototype theatre; permission-to-test delays; evidence not legible to buyers/regulators.

Stage	Primary objective	Key decisions	Typical artifacts	Failure mode
Handoff / scaling trigger	Transfer to an owner who can deploy, procure, regulate, or scale.	External owner decides to adopt; PD/PM ensures the decision package is complete and timed to their process.	Handoff package: evidence bundle; implementation playbook; procurement dossier; IP/data agreements.	No owner or budget on the other side; results stranded; follow-on becomes 'someone else's problem'.

Note: Agencies differ in emphasis. IARPA institutionalises independent test and evaluation; ARPA-E embeds commercialisation and techno-economic evidence early; DARPA and SPRIND often use staged competitions/challenges to create repeatable 'truth moments'.

Minimum decision cadence (across agencies)

- **Programme launch:** single accountable decision-maker approves start based on a clear charter (avoid multi-body sign-off).
- **Execution:** predictable milestone reviews (often quarterly) where the PD/PM can stop, redirect, or reallocate funds.
- **Portfolio reviews:** leadership reviews the portfolio (health, balance, transitions) without taking over technical decisions.
- **Transition hinge points:** explicit moments when adopters/buyers/regulators confirm evidence requirements and ability to act.

3.4 Instrument mix and contracting (incl. OTs; IP and data rights)

Cross-agency pattern

- **Instrument selection is a design lever:** ARPA agencies choose mechanisms that fit the performer (startup vs. university vs. incumbent), the stage (exploration vs. prototyping vs. scaling), and the desired level of control (milestone enforceability, data rights).
- **Milestone-based governance is common across instruments:** even with assistance awards, agencies embed performance milestones and retain the ability to stop or redirect work.
- **IP and data rights are treated as transition variables.** Agencies often seek enough rights to ensure government/public benefit and follow-on use, while avoiding terms that deter top performers.

Meaningful variants across agencies

- **OT-centric flexibility:** ARPA-H is comparatively OT-centric and uses negotiable terms to tailor IP, data, and milestone payment structures; DARPA uses OTs heavily for prototypes and to attract non-traditional performers.
- **Assistance-centric model:** ARPA-E uses cooperative agreements (financial assistance) and compensates by building strong award governance and commercialisation support.
- **FAR-centric procurement contracts:** IARPA typically uses FAR-based R&D contracts, with defaults that favour government rights in data/software unless restricted rights are asserted.
- **EU legal context:** SPRIND uses equity and challenge instruments but applies guardrails (e.g., co-investment requirements; government licensing rights) to manage state-aid compliance. Procurement pathways such as PCP/PPI are more central in Europe than in the U.S. ARPAs

Design questions for NADI

- Which instruments should NADI treat as 'core' vs 'optional' on day one (subsidies, contracts, PCP/PPI, prizes, equity/quasi-equity)?
- How will NADI preserve speed in contracting while meeting Dutch/EU procurement requirements (standardised templates; delegated authority; two-step intake; stage-gated contracting)?
- What IP/data posture best supports transition in NADI's target domains (license-back for public use; data sharing requirements; affordability/access conditions where relevant)?
- If NADI wants an equity lane, what compliance architecture is required (state-aid analysis, co-investment rules, governance for managing stakes/returns)?

3.5 Organisation and talent (PD/PM model)

Cross-agency pattern

- **Term-limited PDs/PMs are the keystone.** Agencies recruit domain leaders for fixed tours to bring fresh ideas and urgency and to avoid sclerotic empires.
- **PDs/PMs are empowered to act as 'general managers' of programmes:** they set vision, assemble performer portfolios, manage milestones, and orchestrate transition pathways.
- **A small core is wrapped with enabling support** (contracting/legal, finance, communications, T&E, transition/commercialisation) to compensate for variability in PD/PM management skill and to keep cycle time short.

Meaningful variants across agencies

- **Clearance:** IARPA often requires security clearances, narrowing the available talent market.
- **Compensation and hiring flexibilities:** SPRIND faced early constraints from German public pay rules; the Freedom Act expanded the ability to hire and set salaries. ARIA was designed with hiring flexibility from inception.
- **Embedded transition teams:** ARPA-E institutionalises T2M as a standing capability; other agencies rely more on ad hoc transition partners or programme-specific structures.

Design questions for NADI

- How will NADI recruit PDs/PMs: open calls, active headhunting, or a hybrid? What selection criteria demonstrate the ability to run an ARPA programme (not just scientific excellence)?
- What term limits, conflict-of-interest rules, and post-service restrictions are required to preserve integrity while remaining attractive to top talent?
- What enabling functions must be centralised (e.g., contracting/legal) versus embedded in programmes (e.g., transition leads), and what minimum staffing ratios are required?

3.6 Pipelines and adoption routes (transition mechanisms; buyer involvement)

Cross-agency pattern

- **Transition is an explicit workstream with artefacts and gates:** identifying end users/buyers early, mapping the adoption pathway, and testing 'plausible pathway' assumptions is core to programme design.
- **Agencies frequently rely on partners for scale:** they can de-risk and prototype, but downstream procurement, deployment, regulation, and financing often are external.

- **Buyer involvement is a strong predictor of transition success** when the buyer is government (defence/intel) or when regulation and infrastructure dominate (energy/health).

Meaningful variants across agencies

- **Defence/intelligence:** DARPA and IARPA commonly use transition partners (services/agencies) who participate in reviews and absorb outputs; adoption is constrained by downstream acquisition or operational priorities.
- **Energy:** ARPA-E emphasises market maps, commercialisation milestones, and a dedicated T2M function; downstream adoption depends on industry, utilities, and capital markets.
- **Health:** ARPA-H must integrate regulators, payers, providers, and patient/community pathways; IP and data sharing can be used to shape incentives for adoption.
- **Europe:** SPRIND blends transition routes — venture creation and market scaling, challenge structures, and (where relevant) public procurement pathways.

Design questions for NADI

- For each NADI programme type, what is the intended 'handoff point' (prototype, validated evidence, demonstrator, early deployment)?
- Which partners must be 'in the room' from day one for credible transition (public procurers, regulators, standards bodies, investors, corporates)?
- What minimum commitments should be required from adopters/buyers before launch (letters of intent, co-funding, data access, testbed availability)?
- Does NADI need a dedicated transition function (ARPA-E-like) or can it rely on programme-specific transition partners? What are the risks of each?

3.7 Domain differences: how models shift across defence, energy, health, and intelligence

Although the six agencies share a recognisable ARPA operating logic, the 'field' (defence, intelligence, energy, health) materially shapes how that logic is operationalised. The key driver is market/adopter structure: who is the eventual 'customer', who bears adoption costs, and what evidence/regulatory gates determine whether a capability can be deployed at scale.

The table below summarises domain-specific patterns that repeatedly appear in our benchmark and are especially relevant to Dutch ministries that are considering where (and how) NADI should engage.

Domain archetype (example agency)	Market & adopter structure	What 'transition' means (and what is evidence)	Implications for the ARPA operating model
Defense (DARPA)	Typically, a single or small set of public buyers; mission urgency can justify high risk. Adoption often requires a committed sponsor inside the armed forces/MoD who can carry a capability into a programme of record.	Transition often means insertion into defence capability development and acquisition. Evidence packages emphasise T&E, performance in representative environments, reliability/safety, and integration with existing systems.	PMs need operator/acquisition context, as well as access to test infrastructure. Programmes benefit from early sponsor engagement and clear handoff points; the agency can prototype quickly, but scaling often depends on external acquisition authorities.
Intelligence (IARPA)	Sensitive operational environments; 'customer' is an Intelligence Community (IC) partner rather than a commercial market. Data access, security constraints, and validation are central to the process.	Transition often means operational adoption by an IC customer. Evidence stresses 'technical truth': independent testing, reproducibility, and performance on mission-representative datasets; publication is constrained, but evaluation remains rigorous.	Independent T&E is more structurally embedded; PMs operate with tighter security/compliance envelopes. Transition partners are typically engaged from the outset to ensure deliverables can be operationalised.
Energy & climate (ARPA-E)	Diffuse private markets with heavy regulation and infrastructure lock-in (utilities, permitting, standards). Adoption often depends on economics, bankability, and ecosystem readiness (suppliers, offtakers, project finance).	Transition often means first-of-a-kind pilots/demonstrations, followed by commercial-scale up. Evidence packages emphasise techno-economic analysis, manufacturability, supply chains, permitting, and customer/offtaker commitments.	T2M functions become core, not optional. Programmes may require heavier engagement with incumbents and financiers; 'scale' is frequently gated by non-technical constraints (permitting, interconnection, standards).
Health & biomedical (ARPA-H)	Multi-actor system (patients, providers, payers, regulators, manufacturers). Incentives are fragmented; safety/ethics constraints are high; time-to-impact can be long.	Transition can mean multiple things: regulatory approval, reimbursement coverage, clinical adoption, or public-health deployment. Evidence packages may require clinical validation, regulatory strategy, and pathways for real-world integration.	Programmes often need regulatory, clinical operations, and health-system integration expertise. Partnering with delivery systems (hospitals/health services) can be as important as partnering with companies.
Cross-domain (ARIA, SPRIND)	Portfolio spans multiple sectors; transition pathways vary by programme. Cross-domain agencies must avoid assuming a single 'default' customer or scaling route.	Transition definitions are programme-specific (public procurement, commercialisation, regulation, standards, open-source, etc.). Therefore, evidence packages vary and should be tailored to the chosen pathway.	Requires deliberate 'lane' design: shared ARPA core processes plus domain-specific transition playbooks/support. If not managed explicitly, cross-domain scope can dilute focus and overwhelm a small core team.

Design questions for NADI

- If NADI spans multiple domains, should it treat 'transition' as domain-specific and build explicit transition playbooks (and partner commitments) per programme — rather than relying on a single default handoff model?
- Should programme design align the evidence package to the relevant gate: operational testing and evaluation for defence/intelligence; techno-economics and bankability for energy; regulatory/clinical integration for health?
- Can support functions be materially different by lane: e.g., acquisition/procurement expertise and test infrastructure access (defence); security and independent evaluation (intelligence); finance/offtaker engagement (energy); clinical/regulatory operations and health-system partnerships (health)?
- If NADI spans multiple domains, how pertinent is the potential failure mode of cross-domain agencies where there is 'one operating system, many markets': programmes are run well, but the adoption pathway is under-specified or mismatched to the domain's real decision-makers?

3.8 Comparative positioning (interaction with national innovation instruments)

Cross-agency pattern

- **ARPA agencies typically occupy a 'white space'** between basic research funding and mission-oriented deployment instruments. Their comparative advantage is coordinated, time-bound, milestone-driven bets — not broad capacity funding.
- **They depend on the wider innovation system** for both inputs (science base, talent, startups, industry) and outputs (scale financing, procurement programmes, regulatory pathways).
- **Clear positioning reduces duplication and helps secure political legitimacy:** agencies justify why they are needed in addition to existing instruments.

Illustrative positioning differences

- DARPA complements service laboratories, defence acquisition, and other DoD innovation entities; it is not designed to own production procurement.
- ARPA-E complements DOE's Office of Science and applied energy programmes; it often hands off to private capital, industrial partners, or other deployment finance mechanisms.
- ARPA-H complements NIH institutes and other HHS agencies; it aims to break through translational barriers that do not fit within conventional grant or procurement mechanisms.
- ARIA was explicitly created to complement (not replace) UKRI by funding bets that are too risky or unconventional for standard peer review.
- SPRIND complements Germany's existing innovation funding and industrial policy tools by taking leap-style bets and, increasingly, by using investment-like instruments within legal constraints.

Design questions for NADI

- Which gaps in the Dutch innovation system are NADI uniquely positioned to fill (coordination failures, high uncertainty, absence of a natural 'owner')?
- How will NADI interact with existing Dutch instruments (innovation subsidies, mission-driven programmes, procurement tools) without becoming another layer of grant funding?
- What 'handoff' relationships must be built early (e.g., procurement authorities, deployment funds, regulators) so NADI outputs have a path to scale?

3.9 Budget and scale (typical budgets; staffing ratios)

Across ARPA-type agencies, a consistent pattern is a lean core staff relative to programme budgets, enabled by the use of external performers and, in some cases, contractors for technical and administrative support. However, public reporting practices vary, and comparisons are difficult.

Indicative scale (publicly reported or commonly cited figures; not fully harmonised)

Agency	Budget scale	Staff scale	Notes on comparability
DARPA	In the billions (annual)	~200–220 government employees; ~100 PMs	Budgets vary by fiscal year and are reported through DoD RDT&E lines; staff numbers often cited include government civilians, plus contractor support.
ARPA-E	Hundreds of millions (annual)	~100 staff	Assistance awards dominate; DOE reporting provides programme budget context; staffing is relatively stable.
ARPA-H	Separate HHS budget line (annual)	Statutory cap 210; early staffing ~100+	Early years still ramping; some budget detail in annual justifications; staffing levels changing.
IARPA	A few hundred million (annual)	~100 staff	Some programmes classified; public figures are approximate; typical programme sizes reported in public materials.
ARIA	£800m over 4 years (founding settlement)	Lean core (tens to low hundreds)	New agency; public staffing and spend patterns will evolve during ramp-up.
SPRIND	~€1bn over 10 years (indicative); annual allocations vary	Lean core (tens to low hundreds)	Investment-like instruments and challenges complicate comparisons with grant/contract models.

Design questions for NADI

- What is the minimum viable scale for NADI to be credible (PD/PM headcount; enabling functions; programme budget per year) while preserving focus?
- What staff-to-budget ratios are realistic given Dutch contracting and compliance requirements? Where can external support safely substitute for internal headcount?
- How will NADI manage multi-year commitments and budget volatility (carry-over authority; portfolio rebalancing rules)?

3.10 Legal and state-aid compliance (subsidies, procurement, equity)

Legal frameworks are not a peripheral concern: they materially shape which instruments can be used, how quickly contracting can occur, and how investment-like tools can be deployed. The EU context — particularly state-aid and procurement law — creates both constraints and design opportunities (e.g., PCP/PPI pathways).

Cross-agency pattern

- U.S. ARPA operations are governed by U.S. federal acquisition and assistance regimes (FAR and agency supplements), with additional flexibilities (notably OTs).

- **European ARPA-like models must reconcile agility with procurement transparency and state-aid rules**; the most relevant 'proofs' come from SPRIND and, to some degree, ARIA's bespoke legal treatment.

EU-relevant lessons (SPRIND as reference case)

- **State-aid alignment through structuring:** examples include requiring private co-investment when taking equity-like positions and ensuring that the public retains a license to use outcomes even when IP remains with the inventors.
- **Autonomy may require legislative change:** early SPRIND constraints (e.g., public pay limitations, ministry approvals) materially reduced agility; later reforms expanded freedoms while retaining audit/oversight.
- **Procurement vs subsidy boundary management:** when public buyers are involved, PCP/PPI and innovation partnerships can provide compliant pathways for pre-commercial development and early deployment.

Design questions for NADI

- Which legal pathways will NADI use for commissioning work (subsidies, procurement, mixed), and what are the implications for speed and control?
- If NADI uses equity/quasi-equity, what state-aid compliance model will govern terms (co-investment, pricing, governance, return recycling)?
- What standardised legal templates and review workflows are required to keep cycle time short while staying compliant?

4. Cross-cutting design choices for NADI

The benchmark suggests that "ARPA-ness" is created by a small number of interlocking design choices. These choices are not binary: agencies make different trade-offs depending on mission, legal context, and ecosystem. For NADI, making these choices explicit will reduce the risk of creating an agency that has ARPA branding but conventional operating constraints.

4.1 Governance shape and legal form

Takeaways

- **Flat technical governance is a defining feature of ARPA-type agencies.** The main question is how to preserve this while satisfying Dutch accountability norms.
- **Practical test: how many layers sit between PDs/PMs and the Director?** Most ARPA models keep this to one or two.
- **Committee risk:** advisory committees can add legitimacy and domain insight, but if they acquire approval rights over programme execution, the operating model drifts toward conventional grant-making.

Design questions for NADI

- What decisions are reserved to NADI's Director/Board (e.g., programme start approvals, budget envelope, risk policy), and what decisions must be delegated to PDs/PMs (portfolio composition, project terminations, reallocations)?
- What is the minimum set of reporting and evaluation requirements that provide legitimacy without forcing front-loaded approvals?
- What governance arrangements protect NADI from "theme drift" and political micromanagement while retaining democratic accountability?

4.2 PD/PM autonomy and talent model

Takeaways

- **PD/PM quality and autonomy are the keystone.** Agencies recruit leaders who can define a programme vision, manage multiple performers, and make hard stop/redirect decisions.
- **Term limits are a structural lever for urgency and against empire-building**; they also shape incentives and recruitment pipelines.
- **Recruitment is often closer to executive search than to open calls**; open calls can complement but rarely substitute for active scouting.

Design questions for NADI

- Will NADI rely on open recruitment, active headhunting, or a hybrid? Who owns scouting, and how is quality assured?
- What tenure is long enough to deliver outcomes but short enough to preserve dynamism (e.g., 3–5 years as a common reference range)?
- What decision rights are PDs/PMs granted once a programme is approved — especially on kill/scale and instrument selection?

4.3 Instrument toolkit and contracting posture

Takeaways

- **Instrument choice should follow programme logic and the transition environment.** A small number of well-understood tools implemented well is often better than a broad menu implemented slowly.
- **Fast lanes:** Speed is frequently achieved through standardised templates, delegated authority, and multi-step intake (short concept first, then negotiation).
- **Tradeoffs:** In the EU context, a core question is how to combine subsidies, procurement, and investment-like tools while staying state-aid compliant.

Design questions for NADI

- Which instruments are mandatory day-one capabilities (and what is the minimum contracting/legal capacity required)?
- Does NADI need a dedicated contracting unit, or can it rely on a central procurement body while retaining PD/PM control over terms and milestones?
- What IP/data rights policy supports transition and public benefit, and when should it be negotiable vs standardised?

4.4 Transition architecture and public-sector integration

Takeaways

- **Transition has 'hinge points' that must be designed:** programme formation, award negotiation, execution reviews, and post-award scale mechanisms.
- **When the government is a buyer, transition requires procurement pathways and early buyer commitment.** When markets are buyers, transition depends on capital, regulation, and standards; agencies can still shape these, but do not control them.
- **NADI's credibility will depend on how it integrates with public-sector actors** (procurers, regulators) without becoming captive to their slower processes.

Design questions for NADI

- For each programme type, who is the 'buyer' or adopter, and what is the handoff point?
- What minimum transition artefacts are required at launch (transition plan, buyer map, regulatory pathway, data strategy)?
- Does NADI need a standing T2M / Transition function (ARPA-E-like), and if so, where does it sit (central vs programme-embedded)?

4.5 Culture and norms (engineered, not aspirational)

Takeaways

- **ARPA cultures are engineered with structure:** autonomy, term limits, high expectations, and the ability to stop work quickly.
- **Constructive dissent mechanisms matter:** red teams, external reviewers, or internal 'challenge' roles can prevent groupthink while preserving speed.
- **Narratives and legitimacy:** agencies often invest in explaining why failure is acceptable at the project level and how learning is captured.

5. Implementation considerations (12-36 month build)

International experience suggests that early design choices create path dependency. New ARPA-like agencies often face a ramp-up challenge: they are expected to deliver early wins while simultaneously building novel operating capabilities (PD/PM recruiting, contracting, transition). A staged implementation can reduce risk while protecting the core ARPA operating system.

5.1 First 0–6 months: establish the operating system

- **Codify decision rights and guardrails:** define what PDs/PMs can decide, what requires Director/Board approval, and what is non-negotiable (mission filters, ethics, conflicts of interest).
- **Stand up enabling functions early:** contracting/legal, finance, HR (term-limited hiring), and a minimal transition capability.
- **Design PD/PM recruiting as an active process:** build a target list, run structured selection (programme pitch + management simulation), and prepare onboarding materials.
- **Develop first-draft standardised templates:** programme charter, milestone plan, transition plan, contracting templates (aligned to Dutch/EU law).

5.2 Months 6–18: run pilot programmes and learn

- **Launch a small number of programmes (e.g., 1–3) to validate the model before scaling**; treat these as learning vehicles for contracting cadence and kill/scale discipline.
- **Establish independent validation approaches appropriate to domains** (e.g., measurement and evaluation partners; testbeds).
- **Institutionalise review cadence:** quarterly portfolio reviews and clear criteria for downselect, termination, and scale.
- **Build external legitimacy:** publish programme rationales, selection processes, and (where feasible) outcomes and learning.

5.3 Months 18–36: scale selectively and harden interfaces

- **Scale PD/PM headcount and programme volume** only when enabling capacity and decision cadence remain robust.
- **Harden transition interfaces:** formalise relationships with procurement bodies, deployment funds, regulators, and standards organisations; clarify what NADI owns vs. what partners own.
- **Iterate governance based on evidence:** adjust guardrails and templates, but preserve PD/PM autonomy and time-bound programme logic.

5.4 Potential early failure modes to actively prevent

- **Committee governance creep:** advisory bodies evolving into approval gates for programme execution.
- **Underinvestment in enabling functions:** expecting PDs/PMs to 'do everything' without contracting/legal/transition support, resulting in slow cycle time, risk aversion, and/or activities being 'left for later' (creating future issues).
- **Theme drift:** expanding mission boundaries to accommodate stakeholder demands, reducing selectivity and diluting impact.
- **Transition as afterthought:** launching programmes without credible adopters/buyers or a plausible pathway to real-world use.

5.5 Illustrative launches: ARIA & SPRIND

A practical question is how to get from 'green light' to a functioning ARPA-like agency — especially given sequencing paradoxes (e.g., independence is needed to hire innovative leadership, but civil servants are often needed to stand up a public entity, set up controls, and secure budget authority).

ARIA (UK) and SPRIND (Germany) are useful recent European comparators. They differ in legal form and context, but both show that launch is a multi-stage process shaped by a small number of early design decisions about delegation, hiring flexibilities, and how 'temporary' set-up arrangements transition into a durable operating system.

In both cases, the 'design choices' were not made by parliament alone. A small set of actors inside government (the sponsor ministry and finance/central government counterparts) shaped operations, freedoms, delegations, and accountability during the set-up phase; legislation then set outer bounds; and the appointed leadership/board translated these parameters into the day-to-day operating system.

Launch element	ARIA	SPRIND	Practical takeaway
Pre-launch design	Policy design and legislation were prepared by the sponsoring department with central-government involvement; leadership was appointed on a 'designate' basis before the agency legally existed.	A federal innovation commission recommended creation (mid-2019); a new federal company (GmbH) was incorporated at the end of 2019 with ministries as shareholders.	Expect a 'shadow phase' in which core design choices are made within government before NADI formally comes into existence. Make the decision-making locus explicit (who decides what, when).
Legal form / sponsor relationship	Established as a statutory body at arm's length from government; designed with significant operational freedoms but still publicly funded.	Established as a state-owned company with public shareholders; early years highlighted constraints typical of public entities, later addressed through the SPRIND Freedom Act (2023).	Evaluate legal form not only for steady-state governance, but for what it enables during year-1 set-up (hiring, contracting, budget flexibility, risk management).
Leadership sequencing	The CEO and Chair were appointed first; they joined as CEO/Chair-designate prior to legal establishment and could shape early operating decisions before programmes launched.	Founding leadership (Managing Director) was appointed early; the supervisory board was constituted thereafter to provide oversight and legitimacy.	Prioritise appointing the top leadership team early (even as 'designate'). Otherwise, structural choices can become path-dependent and hard to unwind by the subsequent leadership.
Initial governance architecture	Board established with Chair and non-executive members; governance designed to enable programme-level autonomy while meeting public accountability expectations.	Supervisory board inaugurated (2020) to oversee strategy, risk, and selection mechanisms; governance evolved alongside instrument reforms.	Define the smallest governance structure that can: (i) delegate real authority to programme leadership; (ii) satisfy ministerial accountability; (iii) avoid committee micromanagement.
Early staffing approach	Started with a lean core and relied on contracted services for some corporate functions while building internal capability; the first cohort of PDs joined once the core system was in place.	Built an initial team to run Challenges and venture-style instruments; experienced constraints typical of public entities (e.g., hiring and contracting frictions), which prompted later reforms.	Plan a phased staffing ramp: set-up staff + small permanent core, followed by a first cohort of programme leaders once delegations, pay bands, and contracting pathways are clear.

Launch element	ARIA	SPRIND	Practical takeaway
First programme portfolio	Early focus on defining opportunity spaces and recruiting PDs; programmes were developed after governance and core processes were established.	Early visible activity through Challenges and project companies helped build legitimacy; programmes served as 'proof of operating model' while reforms were pursued.	Consider a 'first tranche' of programmes that are feasible within initial constraints but still demonstrate ARPA distinctiveness (clear mission, empowered PDs/PMs, fast iteration).
Managing independence vs. public control	Arm's-length status paired with accountability mechanisms (budget reporting, board oversight, sponsor-department relationship).	As a public entity, SPRIND navigated tight controls; reforms expanded autonomy and instrument flexibility along with expectations for evaluation.	Independence is not binary. Use explicit delegations, transparency choices, and evaluation plans to earn trust while preserving speed and PD/PM autonomy.
Instrument flexibility	Designed to use a range of contracting mechanisms; operational detail had to be built (templates, legal review patterns, risk posture).	Initially more constrained; Freedom Act expanded scope for action and funding (incl. instruments and multi-year flexibility).	Instrument 'availability' depends on implementation detail. Prioritise 2–3 ready-to-use pathways on day one, then expand.
Learning and adaptation	First year spent building the operating system and opportunity pipeline; subsequent years expand programme throughput.	Evaluation and reforms indicate an iterative approach: start, learn, then adjust the legal/operational framework.	Treat launch as iterative: set evaluation moments (6/12/24 months) and pre-identify which rules/controls are candidates for adjustment.

Selected launch milestones

- ARIA (UK): Chair and CEO appointed (July 2022); ARIA legally established (January 2023); first cohort of PDs joined (from October 2023); initial programmes developed and launched across a small number of opportunity spaces (2023–2024).
- SPRIND (Germany): Federal innovation commission recommends creation (mid-2019); SprinD GmbH founded (December 2019); supervisory board inaugurated (September 2020); Freedom Act expands autonomy/instrument flexibility (end-2023); first challenge winners (2024).

A concrete way to address the sequencing paradox is to treat 'set-up' as a time-boxed joint venture between a small civil-service implementation team and the incoming NADI leadership (designate), with explicit decision packages. Those packages typically include: (i) the minimum viable instrument toolkit; (ii) HR/pay and conflict-of-interest rules for programme leaders; (iii) procurement/legal templates; (iv) financial controls calibrated to rapid iteration; and (v) transition interfaces with the relevant ministries and procurement bodies.

Launch-phase questions NADI can answer up-front (to reduce path dependence):

- Who inside the government owns 'design authority' during set-up (sponsor ministry, inter-ministerial steering group, finance ministry), and which decisions are explicitly delegated to the CEO/Chair-designate?
- What is the minimum set of hiring flexibilities needed to recruit high-calibre PDs/PMs, and how will those flexibilities be justified to audit and parliamentary stakeholders?
- Which 2–3 contracting/instrument pathways must be operational on day one (templates, delegated signatures, legal review time-boxes) to avoid a 'slow start'?
- What is the initial transition model by domain (defence procurement, energy demonstrations, health-system integration), and who is accountable for the first customer/partner commitments?

6. Conclusion

Across the six agencies benchmarked, the clearest lesson is that "ARPA-like" is less a sector label than a commitment to an operating model: a small, technically led organisation that repeatedly converts ambitious missions into time-bound, milestone-driven portfolios, using flexible tools, and engineering adoption conditions early. Defence, energy, health, and intelligence differ materially in their market structures and adoption gates — but these differences mostly shape how transition is engineered (and what counts as convincing evidence for adoption), not whether the underlying operating logic applies.

For NADI, the design challenge is therefore to assemble a coherent operating system by making a small set of interdependent choices explicit. The benchmark does not point to a single best model, but it does suggest that some choices are make-or-break: if they are left ambiguous, NADI will tend to revert to the default behaviours of conventional public R&D funding.

A small number of make-or-break choices (and what they imply in practice)

1. Mission selectivity and boundary rules ("right to play")
    - In practice: NADI's mandate is translated into explicit inclusion/exclusion filters that prevent portfolio sprawl (e.g., not incremental R&D; not topics already well served by existing instruments; not problems without plausible transition routes).
    - If under-specified: programmes drift toward breadth, stakeholder appeasement, and incrementalism — reducing the likelihood of true disruptive bets.

2. Delegated authority, legal form, and governance shape
    - In practice: there is a single clear yes/no authority for launching programmes, followed by high autonomy within the approved envelope; advisory bodies provide strategic guardrails rather than technical co-governance.
    - If diluted: committee governance creep and multi-signature approvals slow cycle time, narrow the performer base, and make it harder to take (and actively manage) technical risk.

3. PD/PM model and portfolio decision cadence
    - In practice: term-limited PDs/PMs are recruited through active scouting, are explicitly authorised to stop/redirect/reallocate, and operate with a measurable milestone cadence (with independent evaluation where the domain warrants it).
    - If weakened: kill/scale becomes culturally or politically difficult, portfolios become static grant books, and learning cycles slow down.

4. Tooling and contracting posture
    - In practice: NADI has access to a small set of fast tools (e.g., staged competitions/challenges; milestone-based commissioning; prizes; where feasible, repayable/returnable capital) plus enabling legal/finance support designed to enable speed and experimentation while maintaining accountability.
    - If constrained: NADI struggles to engage non-traditional performers, cannot align tools with programme logic, and becomes dependent on slower external procurement/funding pathways.

5. Transition architecture and integration choreography
    - In practice: each programme selects a transition archetype (public procurement, regulatory/payer pathway, market adoption, or hybrid) and treats transition-grade evidence as a deliverable (demos, validation packages, cost/impact models, reference implementations). Public actors are engaged early enough to keep the transition real, but with clear decision rights that prevent co-management.
    - If deferred: programmes can succeed technically but fail at handoff — producing impressive demonstrations that no buyer, operator, or regulator can act on.

6. Launch sequencing (the "independence vs set-up" paradox)
    - In practice: the first 6–12 months are treated as a deliberate set-up phase with a clear division of labour: civil servants set up the legal/HR/procurement scaffolding and interim controls, while empowered founding leadership rapidly sets the operating norms, hires PD/PM talent, and selects a first wave of pilot programmes.
    - If mishandled: early process lock-in creates long-lived friction (slow hiring; conservative contracting; heavy templates) that is difficult to unwind later — and undermines credibility with the kinds of PDs/PMs and performers NADI aims to attract.

Useful next steps (using this benchmark as a decision tool)

A practical way to apply these findings is to run 2–3 programme-design stress tests against the intended NADI model (for example: one procurement-led, one market-led, and one regulatory-led). For each test, work backwards from the adoption gate: identify who decides, what evidence they require, and what capabilities must exist inside NADI to generate that evidence at speed. This exercise reliably surfaces where legal form, instrument choice, governance layers, and enabling functions are misaligned with the desired operating model — while still being concrete enough to inform near-term implementation planning.

In short, the benchmark supports multiple plausible NADI design directions, but it also makes clear that operational coherence is the binding constraint. Once NADI's ARPA-ness is defined — how selective it is, how much PD/PM autonomy it grants, what tools it can use, and how it engineers transition — the remaining design work becomes aligning legal form, governance, and capability build to those choices.

Renaissance Philanthropy

Evidence Pack: International Benchmark of Coordinated Research Programmes

Produced by Renaissance Philanthropy for Ministry of Economic Affairs, Government of the Netherlands
January 2026

Summary

This Evidence Pack accompanies the main benchmark report. It consolidates: (i) a short methods note; (ii) definitions and a glossary of key terms; (iii) comparative matrices used to synthesise findings across agencies; and (iv) a curated reference library. It is designed to be copy/paste-friendly for internal working documents and to support quick fact-checking of benchmark claims.

Contents at a glance:

A. **Methods note:** approach, validation steps, and data limitations.
B. **Glossary:** key terms, including US/UK/EU procurement and instrument vocabulary.
C. **Core comparative matrices** by benchmark dimension (DARPA, ARPA-E, ARPA-H, IARPA, ARIA, SPRIND) [C1–C9]; **Additional comparative annexes**: field/domain variations; launch/stand-up sequencing; illustrative recent programme exemplars; indicative metrics [C10–C13].
D. **Practical templates and checklists** (draft): reusable artefacts derived from common ARPA operating practices.
E. **Curated reference library:** selected primary sources and evaluations.

Renaissance Philanthropy

A. Methods note (validation steps and limitations)

Approach

We combined structured desk research with iterative synthesis into a common comparison framework. For each agency, we reviewed publicly available primary sources (statutes/mandates, annual reports, official guidance, programme pages), supplemented by evaluations, credible secondary analyses, and expert interviews. Findings were organised into matrices by dimension.

Validation steps

- **Triangulation:** major claims were checked against multiple sources where feasible (e.g., statutes, annual reports, and evaluations).
- **Cross-agency normalisation:** we used consistent categories (mission filters; PM model; governance; instruments; adoption/transition; culture; scale) and explicitly noted where terminology differed (PM vs. PD; procurement vs. assistance).
- **Iterative relevance checks:** findings were pressure-tested against experts, including former agency employees, to ensure completeness, accuracy, and salience.
- **Reference hygiene:** where possible, we prioritise stable, official sources (government websites, agency pages, statutes).

Limitations and data gaps

- Budgets, staffing, and programme counts are not always reported consistently across agencies, and newer agencies (ARIA, ARPA-H, SPRIND) have evolving baselines. Quantitative comparisons should be treated as indicative ranges.
- Some IARPA work is classified, limiting public comparability; we rely on unclassified programme materials and publicly described practices.
- Instrument details (e.g., OT templates, negotiation practices, IP terms) vary significantly by programme and contract; we summarise dominant patterns and publicly documented examples, not a comprehensive instrument manual.
- Programme exemplars in C12 are illustrative (selected to show how specific mechanisms play out in practice, with a bias toward recent programs); they are not a representative sample nor an impact evaluation.
- This pack is not legal advice; legal compliance and state-aid/procurement analysis must be confirmed under Dutch and EU law.

B. Glossary (selected terms)

Operating model and programme management

- **Advanced Research Projects Agency (ARPA) model:** A programme-based public R&D operating model built around empowered Programme Directors/Managers (PMs/PDs) running time-bound, milestone-driven portfolios aimed at high-risk, high-reward outcomes and real-world adoption.
- **Programme Manager/Director (PM/PD):** the individual responsible for designing, launching, and actively managing a programme portfolio (milestones, performer selection, re-scoping, termination, and transition planning).
- **Heilmeier Catechism:** A set of framing questions used (originating at DARPA) to clarify what a programme is trying to do, why it matters, what is new, and how success will be measured.
- **Milestone-driven:** Programme structure where continued funding is conditional on meeting explicit technical (and sometimes transition) milestones; enables fast kill/scale decisions.
- **Kill/scale:** A discipline of stopping underperforming approaches and reallocating resources to stronger ones based on evidence and milestone results.
- **Portfolio logic:** Running multiple technical approaches in parallel early, then down-selecting based on evidence; this helps manage uncertainty and avoid single-bet failure.
- **Technology Readiness Level (TRL):** a commonly used (often 1–9) scale describing maturity from basic principles (low TRL) to deployed systems (high TRL).
- **Test & Evaluation (T&E):** independent validation and measurement activities to assess performance claims and support objective go/no-go decisions.

Instruments, contracting, and procurement

- **Broad Agency Announcement (BAA):** A U.S. solicitation mechanism commonly used for R&D acquisitions; enables evaluation of proposals for scientific/technical merit rather than lowest price.
- **Federal Acquisition Regulation (FAR):** the U.S. government-wide framework governing procurement contracts.
- **Other Transaction (OT):** A U.S. legal instrument that provides flexibility outside standard procurement rules for R&D and prototyping (and, in some cases, follow-on production).
- **Cooperative agreement:** A U.S. financial assistance instrument (like a grant) used when substantial government involvement is expected during execution.
- **Prize challenge:** A mechanism to incentivise results through competition and awards; often used to broaden participation and accelerate iteration.
- **SBIR/STTR:** U.S. programs funding early-stage R&D at small businesses (and partners) through phased awards; used by several agencies as part of an innovation pipeline.
- **Pre-Commercial Procurement (PCP):** An EU procurement approach for buying R&D services in phases (competing suppliers), used to develop solutions not yet available on the market.
- **Public Procurement of Innovative solutions (PPI):** EU procurement of innovative products/services that are near-market but not yet widely deployed; the public sector acts as an early adopter to create demand for scale-up.
- **Innovation Partnership:** An EU procurement procedure that combines development and subsequent purchase within a single contract framework, under specified conditions.

- **State aid:** EU rules governing when public support to firms may distort competition; relevant for grants, subsidies, and equity-like instruments.
- **Returnable capital:** Funding structured so that public funds can be repaid if a project succeeds (e.g., revenue share or repayment triggers); used by some ARPA-like bodies to recycle capital and manage risk.
- **Convertible instrument / SAFE-like structure:** Equity-like instruments that convert to equity under certain conditions (e.g., future financing); design must be aligned to public-sector risk controls and state-aid rules.

Transition, adoption, and ecosystem integration

- **Transition sponsor/partner:** A downstream actor (e.g., a defence service, agency, utility, hospital system, regulator) with authority and budget to adopt, procure, or operationalise the outcome.
- **Tech-to-Market (T2M):** ARPA-E's commercialisation and transition support function, including artefacts and services that pressure-test adoption pathways and support follow-on financing/partnerships.
- **Activation partner:** ARIA term for organisations that help translate programme outputs into adoption or field activity (e.g., testbeds, deployment partners), reducing the burden on small internal teams.
- **Intelligence Community (IC):** A group of separate U.S. federal government intelligence agencies and subordinate organisations that work to conduct intelligence activities which support the foreign policy and national security interests of the United States.
- **Testbed/sandbox:** A controlled environment (technical, operational, or regulatory) for piloting solutions and learning under real-world constraints.

Governance and legal forms (UK/Germany examples)

- **Freedom of Information regimes (FOI/FOIA):** Some ARPA-like bodies have exemptions to protect sensitive programme work, which affects transparency norms and data availability.
- **Gesellschaft mit beschränkter Haftung (GmbH):** a German limited-liability company form. SPRIND is structured as a federally owned GmbH, with public shareholder oversight.

C. Comparative matrices by dimension

The following matrices summarise key comparative observations across agencies. They are intended as working references and do not attempt to capture every policy nuance. Additional depth was captured in our raw diligence bundle and the curated reference library below.

C1. Mandates and governance

How the agency is constituted and governed; where autonomy sits; and what oversight mechanisms exist, without introducing committee approval rights over programme execution.

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Department-embedded (DoD) with strong internal autonomy; flat PM→Office Director→Director chain; mission anchored in defence; ex post oversight via budget/audit; strong internal program-framing norms (e.g., Heilmeier).	https://www.darpa.mil/about https://www.darpa.mil/about/program-managers https://www.darpa.mil/about/heilmeier-catechism
ARPA-E	DOE agency with a statutory mission; portfolio programs designed and led by Programme Directors; Tech-to-Market (T2M) is a standing capability; largely assistance-based but with strong governance.	https://arpa-e.energy.gov/ https://arpa-e.energy.gov/about/arpa-e-at-a-glance https://arpa-e.energy.gov/about/tech-to-market
ARPA-H	HHS/NIH placement; broad tool menu with strong OT emphasis; FOIA applies with statutory protections for confidential commercial/financial information; statutory headcount cap; transition network-building through ARPANET-H.	https://arpa-h.gov/about/faqs https://arpa-h.gov/engage-and-connect/other-transaction-community https://arpa-h.gov/engage-and-connect/arpanet-h
IARPA	ODNI component; mission tied to Intelligence Community needs; governance emphasises rigorous metrics and (often) independent evaluation; procurement contracts are common.	https://www.iarpa.gov/ https://www.iarpa.gov/who-we-are/about-us https://www.iarpa.gov/engage-with-us/open-baas
ARIA	Arm's-length UK body with high autonomy under ARIA Act; ministers cannot direct individual funding; FOI exempt; board + CEO; designed to complement UKRI.	https://www.legislation.gov.uk/ukpga/2022/4/notes/division/6/index.htm https://www.aria.org.uk/how-we-work/ https://www.aria.org.uk/media/dbefok51/aria-annual-report-2024-2025.pdf
SPRIND	Federally owned GmbH with supervisory board; Freedom Act expanded autonomy (incl. hiring/investment flexibilities); EU procurement/state-aid constraints remain central; evaluation and reporting support accountability.	https://www.sprind.org/en/overview https://www.sprind.org/en/corporate-governance https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf

C2. Challenge definition and 'ARPA-worthy' filters

How agencies select and frame problems such that they are high-impact, high-uncertainty, and tractable within time-bound programmes.

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Uses explicit problem-framing disciplines (e.g., Heilmeier questions); PMs develop time-bound programs within mission boundaries; emphasis on technical surprise and demonstrable prototypes.	https://www.darpa.mil/about/heilmeier-catechism https://www.darpa.mil/about/program-managers https://www.darpa.mil/research/opportunities/baa
ARPA-E	Programme Directors define 'white space' problems (too risky for the private sector); screening integrates techno-economic plausibility and commercialisation hypotheses.	https://arpa-e.energy.gov/about/arpa-e-at-a-glance https://arpa-e.energy.gov/about/tech-to-market https://arpa-e.energy.gov/about/tech-to-market/commercialization
ARPA-H	Challenge framing emphasises outsized health outcomes and must include adoption system constraints (regulatory, reimbursement, workflows); uses programme constructs and targeted solicitations.	https://arpa-h.gov/about/faqs https://arpa-h.gov/explore-funding/programs/adapt https://arpa-h.gov/explore-funding/programs/paradigm
IARPA	Problem selection is anchored in the Intelligence Community's needs; the solicitation process often begins with white papers; and success metrics and evaluation plans are central to programme design.	https://www.iarpa.gov/engage-with-us/open-baas https://www.iarpa.gov/funding-opportunities https://iarpa-ideas.gov/client/userguide.pdf
ARIA	Uses opportunity spaces and Programme Director-led theses; bets can be unconventional and exploratory; relies on boundary rules to avoid fragmentation.	https://www.aria.org.uk/how-we-work/ https://www.aria.org.uk/opportunity-spaces https://www.aria.org.uk/about-aria/our-team/program-directors
SPRIND	Uses staged Challenges and venture-building to test disruptive ideas; challenge design relies on downselect and staged funding to manage uncertainty.	https://www.sprind.org/en/actions/challenges/articles/overview https://www.sprind.org/en/actions/challenges/antiviral https://www.sprind.org/en/faq

Renaissance Philanthropy

C3. Programme anatomy

Typical programme structure, duration, TRL focus (where relevant), milestone/phase gates, down-selects (where public), kill/scale mechanics, and implied PM workload.

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Typical programme length ~3–5 years; parallel performer portfolio with down-selects; PMs actively manage milestones and funding reallocations; PM tours are term-limited; portfolio budgets can be large (tens of millions per PM per year, indicative).	https://www.darpa.mil/about/program-managers https://www.darpa.mil/sites/default/files/attachment/2024-11/darpa-2024-afr-final.pdf
ARPA-E	Programs run as curated portfolios of projects; strong milestone governance; T2M artefacts integrated; project size and duration vary by program.	https://arpa-e.energy.gov/about/tech-to-market https://arpa-e.energy.gov/programs-and-initiatives/SCALEUP-program
ARPA-H	Programs and initiatives may use OT agreements with milestone payments; place a strong emphasis on data and adoption constraints; and are still establishing a steady-state cadence.	https://arpa-h.gov/engage-and-connect/other-transaction-community https://arpa-h.gov/engage-and-connect/commercialization
IARPA	Often TRL ~3–5; PMs commonly manage ~2 programs concurrently; typical programme scale reported publicly at ~US$20–50M over 3–5 years (indicative); institutionalised independent evaluation supports go/no-go decisions.	https://www.iarpa.gov/engage-with-us/proposers-days https://www.iarpa.gov/images/pdfs/GSS.pdf
ARIA	Programme Directors run portfolios under a thesis; early public materials indicate PD budgets can be on the order of tens of millions of pounds; it uses activation partners to accelerate experimentation and learning.	https://www.aria.org.uk/about-aria/activation-partners/ https://www.aria.org.uk/media/dbefok51/aria-annual-report-2024-2025.pdf
SPRIND	Challenge programs explicitly stage-gate funding; use sprints and downselect to manage risk and uncertainty; venture-building pathways can extend beyond R&D into early scaling.	https://www.sprind.org/en/actions/challenges/articles/overview https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf

C4. Instrument mix and contracting

Core funding and contracting instruments used to engage performers (contracts, grants/assistance, OT, prizes, equity-like tools), and how flexibility is operationalised.

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Broad instrument mix: R&D contracts, grants/cooperative agreements, and OTs for prototypes; contracting office provides standard proposer terms and OT guidance.	https://www.darpa.mil/research/opportunities/baa https://www.darpa.mil/about/offices/contracts-management/proposer-transactions https://www.darpa.mil/about/offices/contracts-management/proposer-general-terms
ARPA-E	Primarily financial assistance (cooperative agreements) with milestone governance; strong FOA infrastructure and award negotiation guidance; complements with commercialisation support rather than procurement levers.	https://arpa-e-foa.energy.gov/ https://arpa-e.energy.gov/innovator-hub/FOA-resources-and-award-reporting/negotiation-doc https://arpa-e.energy.gov/about/tech-to-market
ARPA-H	OT-centric toolset is emphasised; maintains an OT community with FAQs; solicitations and negotiated agreements tailor milestones, data, and IP to programme needs.	https://arpa-h.gov/engage-and-connect/other-transaction-community https://arpa-h.gov/engage-and-connect/other-transaction-community/faqs https://arpa-h.gov/explore-funding/submission-resources-and-FAQs
IARPA	Typically uses BAAs leading to FAR-based R&D contracts; two-step processes and proposer training materials are common; government rights posture can be strong unless restricted rights are asserted.	https://www.iarpa.gov/funding-opportunities https://www.iarpa.gov/engage-with-us/open-baas https://iarpa-ideas.gov/client/userguide.pdf
ARIA	Broad commissioning freedom (beyond standard grants); procurement flexibilities; FOI exemption changes; transparency incentives; and requirements for alternative legitimacy.	https://www.legislation.gov.uk/ukpga/2022/4/notes/division/6/index.htm https://www.aria.org.uk/how-we-work/ https://www.aria.org.uk/funding-opportunities/
SPRIND	Mix of challenges, grants, and investment-like instruments (equity/quasi-equity) with legal guardrails; IP typically remains with inventors with a free non-exclusive government license; uses staged mechanisms to manage uncertainty.	https://www.sprind.org/en/actions/challenges/articles/overview https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf https://dserver.bundestag.de/btd/20/086/2008677.pdf

C5. Organisation and talent

Internal organisation, talent model, and support functions (e.g., contracting/legal, technical staff, transition support, evaluation) that enable PM autonomy and pace.

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	PM model is central; PMs recruited as temporary leaders with high autonomy; flat structure; strong enabling offices around contracting and commercialisation support.	https://www.darpa.mil/about/program-managers https://www.darpa.mil/careers/program-manager https://www.darpa.mil/about/offices/commercial-strategy
ARPA-E	Programme Directors are term-limited and recruited from outside government; T2M is a dedicated support function; the staffing model enables close award management without heavy bureaucracy.	https://arpa-e.energy.gov/careers/program-directors https://arpa-e.energy.gov/about/tech-to-market https://arpa-e.energy.gov/about/tech-to-market/commercialization
ARPA-H	Building a PM workforce under a statutory headcount cap; uses communities to professionalise OT practice; and emphasises cross-sector talent given translation complexity.	https://arpa-h.gov/about/faqs https://arpa-h.gov/engage-and-connect/other-transaction-community https://arpa-h.gov/engage-and-connect/commercialization
IARPA	Smaller PM cadre; engagement mechanisms (seedlings, proposer days) broaden idea intake; institutionalised evaluation supports truth-seeking and kill/scale discipline.	https://www.iarpa.gov/engage-with-us/seedlings https://www.iarpa.gov/engage-with-us/proposers-days https://www.iarpa.gov/images/pdfs/GSS.pdf
ARIA	Programme Directors recruited to run high-risk portfolios; activation partners provide external capacity and engagement; early staffing and processes are still maturing.	https://www.aria.org.uk/about-aria/our-team/program-directors https://www.aria.org.uk/about-aria/activation-partners/ https://www.aria.org.uk/media/dbefok51/aria-annual-report-2024-2025.pdf
SPRIND	Lean core with external partners; the Freedom Act expanded hiring flexibility; challenge and venture-building require strong programme execution and commercialisation skills.	https://www.sprind.org/en/corporate-governance https://www.sprind.org/en/actions/challenges/articles/overview https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf

C6. Adoption/transition pathways

How programmes are designed for transition: who the adopters are (public buyers vs markets), the timing of engagement, and the transition artefacts and partners used.

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Transition relies on early identification of defence users and sponsors; DARPA supports transition/commercialisation but typically hands off to acquisition programs for fielding and scale.	https://www.darpa.mil/sites/default/files/attachment/2025-07/transition-and-commercialization https://www.darpa.mil/about/offices/commercial-strategy
ARPA-E	T2M is the formal transition mechanism; commercialisation plans and milestones are integrated; SCALEUP and engagement with investors/industry support movement toward deployment financing.	https://arpa-e.energy.gov/about/tech-to-market https://arpa-e.energy.gov/programs-and-initiatives/SCALEUP-program https://arpa-e.energy.gov/about/tech-to-market/commercialization
ARPA-H	Transition levers include commercialisation support and network-building (ARPANET-H); OT terms can shape data/IP to support adoption; regulatory and reimbursement constraints are central.	https://arpa-h.gov/engage-and-connect/commercialization https://arpa-h.gov/engage-and-connect/arpanet-h https://arpa-h.gov/engage-and-connect/other-transaction-community
IARPA	Transition to IC customers; technology protection and sensitive-data constraints matter; rigorous evaluation strengthens credibility for partner-agency adoption decisions.	https://www.iarpa.gov/images/pdfs/GSS.pdf https://www.iarpa.gov/research-programs/research-and-technology-protection
ARIA	Transition expected through ecosystem uptake and entrepreneurship; activation partners and commissioning flexibility support translation; formal transition mechanisms are still emerging.	https://www.aria.org.uk/about-aria/activation-partners/ https://www.aria.org.uk/funding-opportunities/
SPRIND	Blend of venture creation, challenge progression, and (where relevant) public-sector pathways; investment-like tools allow continued support through early scale-up stages, with state-aid guardrails.	https://www.sprind.org/en/actions/challenges/articles/overview https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf

C7. Comparative positioning in the national innovation system

How each agency fits alongside other national instruments, and which functions it does (or does not) attempt to own (deployment programmes, procurement, capital, regulation).

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Operates upstream of DoD acquisition; complements service labs and other DoD innovation entities; not designed to own production procurement.	https://www.darpa.mil/about https://www.darpa.mil/sites/default/files/attachment/2025-07/transition-and-commercialization
ARPA-E	Gap-filler in energy innovation system; complements DOE offices and private capital by funding risky, time-bound bets and translating toward commercialisation.	https://arpa-e.energy.gov/about/arpa-e-at-a-glance https://arpa-e.energy.gov/about/arpa-e-at-a-glance/engagement
ARPA-H	Complement to NIH/HHS mechanisms; aims to tackle translational barriers that do not fit conventional grants or procurement.	https://arpa-h.gov/about/faqs https://arpa-h.gov/engage-and-connect/commercialization
IARPA	R&D engine for the Intelligence Community; complements mission agencies by running coordinated portfolios and evaluations, then transitioning results to operational owners.	https://www.iarpa.gov/who-we-are/about-us https://www.iarpa.gov/engage-with-us/open-baas
ARIA	Created to complement UKRI by funding bets too risky or unconventional for standard peer review and grantmaking.	https://www.legislation.gov.uk/ukpga/2022/4/notes/division/6/index.htm https://www.aria.org.uk/how-we-work/
SPRIND	Complements German innovation instruments by taking leap bets and supporting venture creation and scaling within EU legal constraints.	https://www.sprind.org/en/overview https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf

C8. Budget and scale

Indicative agency scale: budgets, staffing, and programme throughput (not fully harmonised across agencies due to reporting differences).

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Lean core workforce supports a very large external performer base; the budget is in the billions annually (varies by fiscal year and DoD line items).	https://www.darpa.mil/sites/default/files/attachment/2024-11/darpa-2024-afr-final.pdf https://www.darpa.mil/about
ARPA-E	Staff ~100; annual budgets on the order of hundreds of millions; runs multiple programme areas and funding opportunities each year.	https://arpa-e.energy.gov/about/arpa-e-at-a-glance https://arpa-e.energy.gov/
ARPA-H	Statutory headcount cap (210); early staffing ramp-up ongoing; budget reported through HHS appropriations/justifications; portfolio still maturing.	https://arpa-h.gov/about/faqs https://arpa-h.gov/explore-funding/submission-resources-and-FAQs/submission-faqs
IARPA	Smaller than DARPA; public sources often cite budgets in the few-hundred-million range with ~20–30 programs; staffing commonly cited at ~100.	https://www.iarpa.gov/who-we-are/about-us https://www.iarpa.gov/images/pdfs/GSS.pdf
ARIA	Founding settlement of £800m over 4 years; lean core staffing; public annual reporting emerging as agency ramps.	https://www.aria.org.uk/media/dbefok51/aria-annual-report-2024-2025.pdf https://www.aria.org.uk/how-we-work/
SPRIND	Indicative funding on the order of €1bn over a decade; annual allocations vary; publishes evaluation and annual reporting materials.	https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf https://cms.system.sprind.org/uploads/2025_05_20_PCGK_Bericht_2024_final_18e9771ca6.pdf

C9. Legal and state-aid compliance considerations

How legal regimes shape instrument choice, contracting speed, transparency obligations, and the feasibility of equity-like tools—especially relevant for an EU-context agency like NADI.

Agency	Benchmark observations (dimension-specific)	Source pointers (examples)
DARPA	Operates under U.S. procurement and assistance regimes plus OTs for research/prototypes; standardised legal terms and OT guidance support speed; FOIA applies with standard protections.	https://www.darpa.mil/about/offices/contracts-management/proposer-transactions https://www.darpa.mil/about/offices/contracts-management/proposer-general-terms https://www.darpa.mil/research/opportunities/baa
ARPA-E	Primarily financial assistance; compliance follows DOE assistance rules; standardised negotiation and reporting processes support governance and auditability.	https://arpa-e.energy.gov/about/arpa-e-at-a-glance/authorization https://arpa-e.energy.gov/innovator-hub/FOA-resources-and-award-reporting/negotiation-doc
ARPA-H	OT and other tools allow tailored contracting terms; FOIA applies with statutory CCI protections; program-specific negotiation is common.	https://arpa-h.gov/engage-and-connect/other-transaction-community https://arpa-h.gov/engage-and-connect/other-transaction-community/faqs
IARPA	FAR-based contracting is common; data/rights posture can be strong; classified environments add additional compliance constraints and can affect collaboration and evaluation.	https://www.iarpa.gov/engage-with-us/open-baas https://www.iarpa.gov/images/pdfs/GSS.pdf
ARIA	The ARIA Act provides a bespoke regime, including FOI exemptions and procurement flexibilities; autonomy is legally protected, while legitimacy is maintained through alternative reporting and governance.	https://www.legislation.gov.uk/ukpga/2022/4/notes/division/6/index.htm https://www.aria.org.uk/media/dbefok51/aria-annual-report-2024-2025.pdf
SPRIND	EU/German procurement and state-aid rules are central; the Freedom Act expanded flexibilities; investment tools use guardrails (e.g., co-investment) and IP licensing to align with state-aid principles; PCP/PPI pathways are relevant for public buyers.	https://dserver.bundestag.de/btd/20/086/2008677.pdf https://dserver.bundestag.de/btd/20/127/2012726.pdf https://cms.system.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf

C10. Field/domain variations (illustrative)

Across the benchmark set, the ARPA operating model is recognisable, but it plays out differently across fields and domains. Differences in market structure, buyer power, regulatory pathways, and validation requirements influence programme design and transition strategy.

Field/domain	Typical adopter and transition destination	Operating implications for the ARPA-type model
Defence (DARPA archetype)	The primary adopter is the state (defence services) via mission partners; the transition often occurs into acquisition programmes, operational units, or defence primes.	Early 'transition sponsor' identification is central; demonstration in realistic environments; procurement pathways often separate from R&D funding; strong emphasis on T&E, security, and system integration.
Intelligence (IARPA archetype)	The primary adopter is the Intelligence Community; the transition typically involves classified/mission systems and workflows.	Greater sensitivity/classification; evaluation regimes may require controlled datasets and independent testing; adoption depends on IC integration owners; public evidence base thinner—governance and legitimacy rely more on internal oversight.
Energy/climate (ARPA-E archetype)	Adoption often requires commercial markets, regulated utilities, and infrastructure; transitions frequently via pilots/demonstrations, corporate partnerships, and follow-on project finance.	'Valley of death' is often scale/manufacturing and project-finance, not only technical proof; cost-share and industry engagement are common; commercialisation support functions (T2M-like) materially affect outcomes.
Health/biomed (ARPA-H archetype)	Adoption involves complex ecosystems (providers, payers, FDA/regulators, hospital systems, supply chains).	Evidence thresholds can be clinical/regulatory; programmes may need data infrastructure, trial design, and regulatory engagement; milestones often combine technical + clinical adoption criteria; transition may require reimbursement and care-delivery integration.
Cross-cutting (ARIA & SPRIND)	Adoption can be via public buyers (digital identity, critical infrastructure) and/or fast-moving markets (software, AI), depending on the topic.	Where markets move quickly, iteration speed and pilot environments matter; where public trust is central, governance/ethics and transparency mechanisms become part of programme design; procurement/standards can be key levers.

C11. Launch and stand-up sequencing (ARIA and SPRIND)

European ARPA-like organisations illustrate that 'launch' is a multi-stage process: policy design and legislation are necessary but insufficient. Early sequencing choices—particularly around interim civil-service scaffolding, delegated authorities, and day-one corporate services—shape the organisation's ability to recruit innovative leadership and execute at pace.

Sequencing element	ARIA (UK)	SPRIND (Germany)
Policy design and enablement	Created by the Advanced Research and Invention Agency Act 2022; designed as an arm's-length body with a board and CEO, and with specific flexibilities (e.g., around transparency/FOI).	Established as a federally owned GmbH (2019) to enable more flexible operations than a classic ministry unit; subsequently adjusted through the SPRIND Freedom Act (2023) to further loosen legal/financial constraints.
Leadership appointment and setting of mandates	Chair/CEO appointed through government/public appointments; early mandate is translated into a small number of 'opportunity spaces' and programmes chosen by leadership rather than a broad open call.	Managing directors appointed; early mandate shaped by shareholders (federal ministries) and company governance; programme mix developed through innovation challenges and venture-building pathways.
Interim setup ('scaffolding') before full autonomy	The sponsoring department supports the initial setup of core controls (finance, HR, governance) so ARIA can operate as a public body with public money; autonomy increases as internal functions mature.	Early operations reportedly faced constraints typical of public ownership (approvals, pay rules); reliance on existing public-sector processes created a 'speed vs control' tension until reforms clarified delegated authorities.
First operating model choices	Lean internal team; heavy emphasis on programme autonomy; experiments with 'activation partners' and external delivery mechanisms to extend capacity without building a large bureaucracy.	Lean internal team; structured innovation challenges with staged funding and down-selects; combination of grants/procurement-like instruments and (in some cases) venture creation/investment structures.
Legitimacy, governance, and guardrails	Designed to tolerate risk while maintaining legitimacy; for sensitive areas (e.g., climate cooling), published oversight and governance arrangements and independent review mechanisms.	Legitimacy was built through public-mission framing and transparency in challenges; governance reforms (SPRINDFG) aimed to reconcile public-money accountability with faster execution and more flexible instruments.
Early programme launch	Started with a small number of opportunity spaces/programmes rather than attempting broad thematic coverage; programme pipeline expands as team and mechanisms stabilise.	Portfolio built through repeated challenge cycles; staged competitions allow early learning and visible progress, but require strong operational capacity (selection, contracting, monitoring).
Reforms after launch	Still early and evolving (processes, transparency practices, partner mechanisms).	Major reform milestone: SPRIND Freedom Act (Dec 2023) increased flexibility; organisation continues to adapt governance and operations based on lessons learned.

C12. Illustrative recent programmes

The table below provides a small set of recent examples (with a bias toward 2019–2025) that show how specific operating features play out in practice. Examples are illustrative—not exhaustive—and are included to help connect abstract design choices (milestones, competition formats, transition partners, evaluation regimes) to concrete programme mechanisms.

Agency	Programme exemplar	Feature(s) illustrated	Why useful	Primary source(s)
DARPA	US2QC / Quantum Benchmarking Initiative (public down-select: 2025)	Independent evaluation & milestone gating; down-select based on benchmark performance	Shows how DARPA can create objective testing regimes for emerging tech and use phased down-selects to manage uncertainty and focus resources	https://www.darpa.mil/news/2025/us2qc-teams-selected
DARPA	AIxCC: AI Cyber Challenge (2023–2025)	Competition + prizes; external event integration (DEF CON); cross-agency collaboration	Shows a 'race' format with public milestones and a large ecosystem, blending R&D with adoption incentives and visible progress signals	https://www.darpa.mil/research/programs/ai-cyber
ARPA-E	SCALEUP / SCALEUP Ready (2019–)	Transition support beyond R&D; scale-up pathway for promising portfolio projects	Illustrates an explicit mechanism for bridging scale/manufacturing/deployment barriers—often the limiting step in energy/climate innovation	https://arpa-e.energy.gov/programs-and-initiatives/SCALEUP-program/SCALEUP-ready
ARPA-E	Tech-to-Market (T2M) function (ongoing)	Embedded commercialisation support; transition artefacts; partner/funder readiness	Shows how a dedicated transition function can be institutionalised (not left to individual PMs), which is especially important when markets and infrastructure are complex	https://arpa-e.energy.gov/about/tech-to-market
ARPA-H	PARADIGM (rural care platform) (2024–)	System-level adoption framing; teaming expectations; real-world deployment constraints	Illustrates how ARPA-H programmes can integrate delivery constraints (rural access, platform logistics) early, not as a late-stage add-on	https://origin.arpa-h.gov/explore-funding/programs/paradigm
ARPA-H	ARPANET-H (health innovation network) (2024–)	Ecosystem network as an adoption lever; hubs/spokes model	Shows an 'infrastructure + network' approach to transition: building pathways and partners that persist beyond individual programmes	https://origin.arpa-h.gov/engage-and-connect/arpanet-h
IARPA	TrojAI (adversarial ML security) (2019–2025)	Evaluation against adversarial conditions; mission-driven AI security	Illustrates how an intelligence-context ARPA uses programme structure to drive measurable capability in a high-stakes, difficult evaluation regime	https://www.iarpa.gov/research-programs/trojai

Agency	Programme exemplar	Feature(s) illustrated	Why useful	Primary source(s)
IARPA	HIATUS (authorship attribution & privacy) (2022–2026)	BAA-based programme with defined duration and transition intent; research security emphasis	Shows a contemporary IARPA programme with explicit timeframe and transition intent, and highlights the role of controlled evaluation artefacts/datasets	https://www.iarpa.gov/research-programs/hiatus
ARIA	Exploring Climate Cooling (2023–)	High-controversy programme governance; oversight committees; legitimacy-by-design	Shows how a small ARPA-like body can build bespoke governance/ethics mechanisms as part of programme design when public trust and risk are central	https://www.aria.org.uk/opportunity-spaces/future-proofing-our-climate-and-weather/exploring-climate-cooling/oversight-and-governance/
ARIA	Scaling Compute (2023–)	Focused, time-bound programme thesis; portfolio of approaches under a PD	Illustrates the ARIA 'programme thesis' approach and how a PD frames a hard, high-uncertainty challenge for portfolio execution	https://www.aria.org.uk/scaling-compute/
SPRIND	Long-Duration Energy Storage Challenge (2022–2025)	Staged challenge format; multiple solution approaches; explicit objective	Shows 'bet on the race, not the horse': using competitive phases to explore diverse approaches, then down-select as evidence accumulates	https://www.sprind.org/en/actions/challenges/energystorage
SPRIND	Broad-Spectrum Antivirals Challenge (2021–)	Challenge structure with proof-of-concept requirements; societal mission with weak private incentives	Illustrates use of a challenge to create a market-shaping innovation pathway where commercial incentives alone are insufficient	https://www.sprind.org/en/actions/challenges/antiviral

C13. Quantitative parameters (indicative; non-harmonised)

Where publicly available, the table below provides indicative quantitative parameters. Values are not fully harmonised (due to different fiscal calendars, reporting practices, and classification limits). For NADI design, treat these as order-of-magnitude signals.

Agency	Indicative annual budget / funding envelope	Indicative staffing / PM scale	Source pointers (examples)
DARPA	FY2024 appropriations ≈ $4.064B; total budgetary resources ≈ $5.098B (USD, in thousands in AFR).	≈250 personnel; nearly 100 programme managers; ≈300 active projects (FY2024 AFR).	DARPA FY2024 Agency Financial Report: https://www.darpa.mil/sites/default/files/attachment/2024-11/darpa-2024-afr-final.pdf
ARPA-E	FY2024 budget request shows total ARPA-E programme direction ≈ $650.2M (request; see DOE budget docs).	Staffing/PM counts are not consistently reported in a single public source; they are typically described as a small, expert team relative to the portfolio.	DOE FY2024 Congressional Budget Request (ARPA-E): https://www.energy.gov/sites/default/files/2023-03/doe-fy-2024-budget-volume-5-v4.pdf
ARPA-H	FY2024 appropriation: $1.5B (ARPA-H budget page).	Early-stage organisation; public staffing metrics are evolving.	ARPA-H budget: https://arpa-h.gov/about/budget
IARPA	Public budget details are limited; funding flows through the Intelligence Community and is not itemised to the same extent as civilian agencies.	Staffing/portfolio details are partially constrained by classification; programme pages and BAAs provide the most consistent public signals.	IARPA programmes: https://www.iarpa.gov/research-programs
ARIA	Multi-year UK funding settlement; annual accounts report Grant-in-Aid and expenditure (not directly comparable to programme budget authority).	50.3 FTE staff at 31 March 2025; average 42.7 FTE across FY2024–25 (annual report).	ARIA Annual Report & Accounts 2024–25: https://assets.publishing.service.gov.uk/media/686f6f07fe1a249e937cbf20/27888_ARIA_AnnualReport_2024-25_Print.pdf
SPRIND	Public envelope is often described as ≈ €1B over 10 years (Germany); annual budget and spending vary and are reported across multiple documents.	Employee count reported in evaluation materials (e.g., ≈80 employees in March 2024, excluding subsidiaries—see evaluation executive summary).	SPRIND Corporate Governance 2024: https://cms.system.sprind.org/uploads/Corporate_Governance_Bericht_SPRIND_Gmb_H_2024_d89330a028.pdf

D. Practical templates and checklists

These templates are not 'requirements'; they are reusable artefacts distilled from common ARPA operating practices. They can help NADI/NADI-design teams translate design choices into concrete decision packages and governance routines.

D1. 'ARPA-worthy' problem filter (Heilmeier-style)

Core framing questions (adapt as needed):

1. What are you trying to do (in simple, concrete terms)?
2. How is it done today, and what are the limits of current practice?
3. What is new in your approach (and why do you think it will work)?
4. Who cares if you succeed (and who is the transition sponsor/adopter)?
5. If you succeed, what difference will it make (and for whom)?
6. What are the key technical risks and the 'unknown unknowns'?
7. What are the measurable mid-course milestones and final 'go/no-go' success criteria?
8. How long will it take and what will it cost (order-of-magnitude)?
9. What would cause you to stop (explicit kill criteria)?

D2. Programme launch decision package (minimum viable)

A programme should typically be launch-ready when the PM/PD can produce:

- A crisp problem statement and 'right-to-play' rationale (why this agency; why now).
- A theory of change/pathway to impact (including transition hypotheses).
- A portfolio plan (multiple approaches, performer types, and an initial down-select logic).
- Milestones, metrics, and an evaluation plan (including who validates).
- Instrument plan (which mechanisms will be used and why).
- A transition/adoption plan (sponsors, pilot sites, procurement pathway hypotheses).
- A resourcing plan (PM time, support needs, contracting/legal, T&E, comms/legitimacy).
- Risk register (technical, adoption, ethics/safety, reputational, legal/compliance).
- A plan for decision cadence (how often evidence is reviewed and who decides).

D3. Kill/redirect/scale cues (portfolio management)

Common cues used by ARPA-type organisations (illustrative):

- Milestones missed without credible new evidence; repeated schedule slippage without learning.
- A competing approach demonstrates superior performance under comparable evaluation.
- Transition sponsor withdraws, or the adoption pathway collapses (unless an alternative is secured).
- Safety/ethics/regulatory constraints become binding in ways that cannot be mitigated within scope.
- Performer capability gaps persist despite remediation; inability to execute at pace.

- Conversely: milestone over-performance + credible transition pull → reallocate resources to scale.

D4. Transition/adoption plan (minimum viable)

At minimum, capture the following early (and update continuously):

- Target adopter(s) and their decision rights (who can say 'yes' to deployment/procurement).
- Adoption constraints (budget cycles, regulatory approvals, integration requirements, standards).
- Pilot/testbed plan (where learning happens; what data is generated; who owns it).
- Procurement/commercial pathway hypothesis (prototype vs scale; who funds follow-on).
- Incentives alignment (why adopters/partners participate; what they get; what they risk).
- Exit/hand-off plan (what 'done' means; who takes ownership after programme end).

E. Curated reference library

This library is selective (non-exhaustive). It prioritises stable primary sources (statutes/mandates, annual reports, official guidance, and programme pages) and a small number of credible evaluations. Where relevant, it includes links to the illustrative programme examples listed.

Cross-cutting and EU/UK procurement context

- Heilmeier Catechism (DARPA): https://www.darpa.mil/work-with-us/heilmeier-catechism
- European Commission: Pre-Commercial Procurement (PCP): https://research-and-innovation.ec.europa.eu/strategy/support-policy-making/support-toolbox/pre-commercial-procurement_en
- European Commission: Public Procurement of Innovative solutions (PPI): https://research-and-innovation.ec.europa.eu/strategy/support-policy-making/support-toolbox/public-procurement-innovative-solutions_en
- European Commission: Innovation Procurement (overview & guidance): https://commission.europa.eu/funding-tenders/find-funding/eu-funding-programmes/funding-opportunities/innovation-procurement_en
- Directive 2014/24/EU (EU public procurement directive; innovation partnership procedure): https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32014L0024
- EU General Block Exemption Regulation (GBER) – consolidated text: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02014R0651-20230701

DARPA

- About DARPA / mission and organisation: https://www.darpa.mil/about-us
- For performers (proposers, solicitations, contracting basics): https://www.darpa.mil/work-with-us/for-performers
- Proposers and submissions guide: https://www.darpa.mil/work-with-us/for-performers/proposers
- Other Transactions (OT) Guide (DARPA): https://www.darpa.mil/work-with-us/other-transactions
- Commercial Strategy Office (transition support function): https://www.darpa.mil/work-with-us/commercial-strategy-office
- Transition & Commercialisation Support Program: https://www.darpa.mil/work-with-us/transition-and-commercialization-support-program
- Transition & Commercialisation Strategy Development Guide (PDF): https://www.darpa.mil/sites/default/files/attachment/2024-03/Transition-and-Commercialization-Strategy-Development-Guide-DARPA.pdf
- Budget and testimony landing page (links to financial reports): https://www.darpa.mil/work-with-us/budgets-and-testimony
- DARPA Agency Financial Report FY2024 (PDF): https://www.darpa.mil/sites/default/files/attachment/2024-11/darpa-2024-afr-final.pdf
- Programme exemplar: US2QC teams selected (news release): https://www.darpa.mil/news/2025/us2qc-teams-selected

- Programme exemplar: AI Cyber Challenge (AIxCC): https://www.darpa.mil/research/programs/ai-cyber

ARPA-E

- ARPA-E about / mission: https://arpa-e.energy.gov/about
- ARPA-E Tech-to-Market (T2M): https://arpa-e.energy.gov/about/tech-to-market
- SCALEUP Ready (programme exemplar): https://arpa-e.energy.gov/programs-and-initiatives/SCALEUP-program/SCALEUP-ready
- ARPA-E annual report library: https://arpa-e.energy.gov/about/annual-reports
- DOE FY2024 Congressional Budget Request (Volume 5; includes ARPA-E): https://www.energy.gov/sites/default/files/2023-03/doe-fy-2024-budget-volume-5-v4.pdf
- ARPA-E Funding Opportunities (FOAs): https://arpa-e-foa.energy.gov

ARPA-H

- ARPA-H about / mission: https://arpa-h.gov/about
- ARPA-H budget: https://arpa-h.gov/about/budget
- Funding opportunities landing page: https://arpa-h.gov/explore-funding
- Programme exemplar: PARADIGM: https://origin.arpa-h.gov/explore-funding/programs/paradigm
- Programme exemplar: ARPANET-H: https://origin.arpa-h.gov/engage-and-connect/arpanet-h
- Programme exemplar: ADAPT: https://origin.arpa-h.gov/explore-funding/programs/adapt
- Other Transaction (OT) community resources: https://arpa-h.gov/resources/other-transaction-community

IARPA

- IARPA overview: https://www.iarpa.gov
- Research programmes (portfolio index): https://www.iarpa.gov/research-programs
- Funding opportunities / BAAs: https://www.iarpa.gov/funding-opportunities
- Programme exemplar: TrojAI: https://www.iarpa.gov/research-programs/trojai
- Programme exemplar: HIATUS: https://www.iarpa.gov/research-programs/hiatus

ARIA

- ARIA homepage: https://www.aria.org.uk
- ARIA 'How we work': https://www.aria.org.uk/how-we-work/
- ARIA Annual Report & Accounts 2024–25 (GOV.UK): https://www.gov.uk/government/publications/arias-annual-report-and-accounts-2024-to-2025
- ARIA Annual Report 2024–25 (print-ready PDF): https://assets.publishing.service.gov.uk/media/686f6f07fe1a249e937cbf20/27888_ARIA_Annual Report_2024-25_Print.pdf
- Advanced Research and Invention Agency Act 2022 (UK legislation): https://www.legislation.gov.uk/ukpga/2022/31/contents/enacted

- Programme exemplar: Exploring Climate Cooling — Oversight & Governance: https://www.aria.org.uk/opportunity-spaces/future-proofing-our-climate-and-weather/exploring-climate-cooling/oversight-and-governance/
- Programme exemplar: Scaling Compute: https://www.aria.org.uk/scaling-compute/

SPRIND

- SPRIND homepage: https://www.sprind.org/en/
- Corporate governance landing page (EN): https://www.sprind.org/en/corporate-governance
- Corporate Governance Report 2024 (PDF): https://cms.system.sprind.org/uploads/Corporate_Governance_Bericht_SPRIND_Gmb_H_2024_d89330a028.pdf
- SPRIND Freedom Act (SPRINDFG) legal text (Bundesgesetzblatt PDF): https://www.recht.bund.de/bgbl/1/2023/415/regelungstext.pdf
- Evaluation summary (Fraunhofer ISI; hosted by SPRIND): https://www.sprind.org/uploads/SPRIND_Evaluation_Zusammenfassung_65119fe433.pdf
- Programme exemplar: Long-Duration Energy Storage Challenge: https://www.sprind.org/en/actions/challenges/energystorage
- Programme exemplar: Broad-Spectrum Antivirals Challenge: https://www.sprind.org/en/actions/challenges/antiviral
- Programme exemplar: EUDI Wallet Prototypes (FUNKE): https://www.sprind.org/en/actions/sovereign-eudi-wallet/eudi-wallet-prototypes-funke