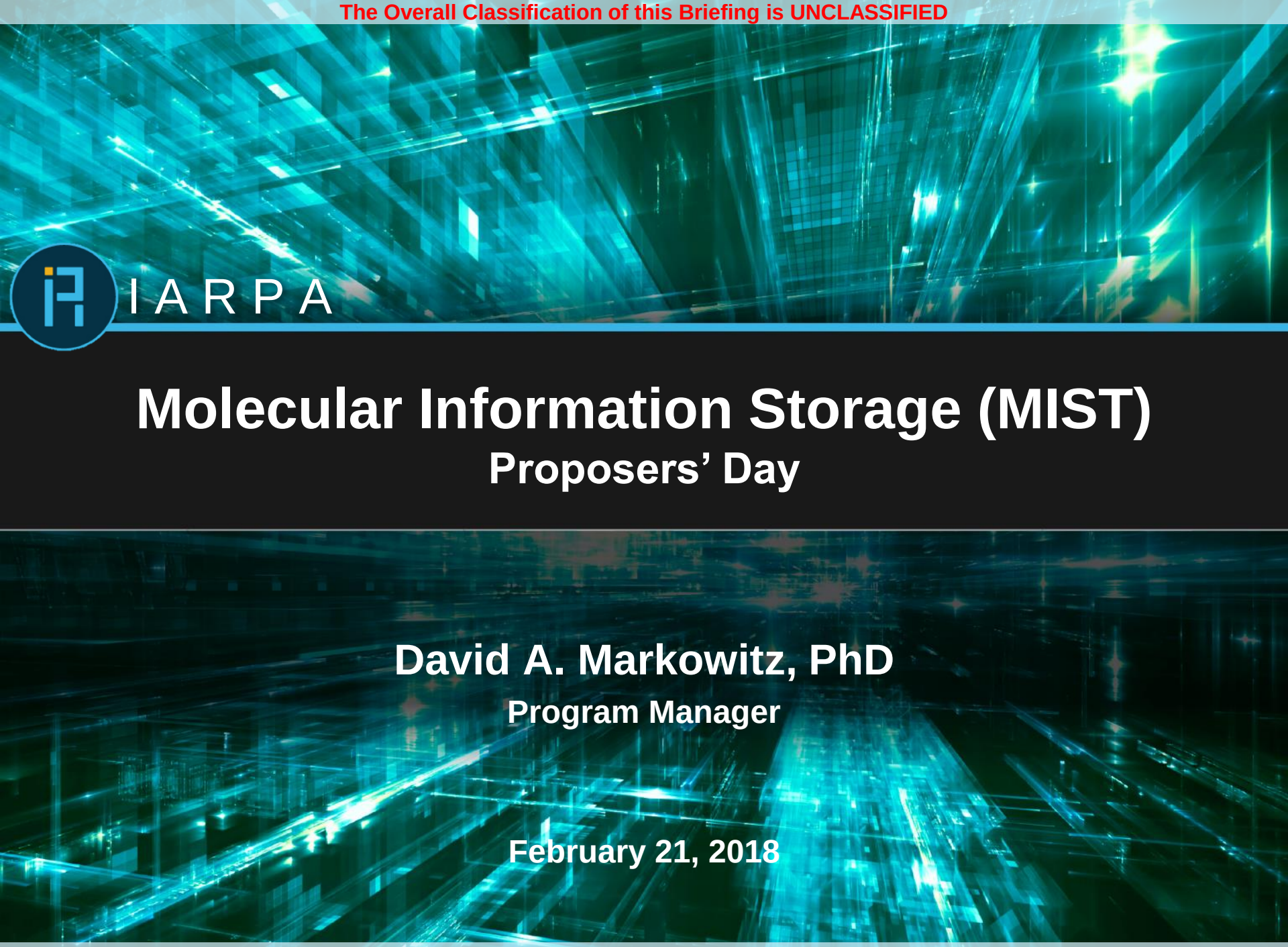


IARPA

Molecular Information Storage (MIST)

Proposers' Day

David A. Markowitz, PhD

Program Manager

February 21, 2018

Disclaimers

- This presentation is provided solely for information and planning purposes
- The Proposers' Day does not constitute a formal solicitation for proposals or proposal abstracts
- Nothing said at Proposers' Day changes the requirements set forth in a BAA
- A BAA supersedes anything presented or said by IARPA at the Proposers' Day

Goals

- Familiarize participants with IARPA's interest in the MIST program
- Please ask questions and provide feedback, as this is your chance to alter the course of events.
- Foster discussion of complementary capabilities among potential program participants, AKA teaming. Take a chance, someone might have a missing piece of your puzzle.

Questions

- During this session, questions should be recorded on note cards. They will be answered for everyone's benefit at a later point in the presentation.

- If/when a BAA is released, questions can only be submitted to the e-mail address provided in the BAA and will only be answered in writing on the program website.

Agenda

Time	Topic	Speaker
9:00am – 9:15am	Welcome, Logistics, Proposers' Day Goals	Dr. David A. Markowitz Program Manager, IARPA
9:15am – 9:45am	IARPA Overview	Dr. William Vanderlinde Chief Scientist, IARPA
9:45am – 10:30am	MIST Program Overview	Dr. David A. Markowitz Program Manager, IARPA
10:30am – 11:00am	Break	
11:00am – 11:20am	Doing Business with IARPA	Acquisition Team
11:20am – 12:00pm	MIST Questions & Answers	Dr. David A. Markowitz Program Manager, IARPA
12:00pm – 1:30pm	No-Host Lunch	
1:30pm – 2:30pm	Offerors' Capabilities Briefings	Attendees (**No Government**)
2:30pm – 5:00pm	Poster Session, Networking and Teaming Discussions	Attendees (**No Government**)

IARPA Overview

Dr. William Vanderlinde, Chief Scientist
Intelligence Advanced Research Projects Activity

Office of the Director of National Intelligence

IARPA

BE THE FUTURE

Office of the Director of National Intelligence
IARPA
BE THE FUTURE

The United States Intelligence Community

IARPA Mission

IARPA envisions and leads *high-risk, high-payoff research* that delivers innovative technology *for future overwhelming intelligence advantage*

- Our problems are **complex** and **multidisciplinary**
- We emphasize **technical excellence** & **technical truth**

IARPA Method

Bring the best minds to bear on our problems

- Full and open competition to the greatest possible extent
- World-class, rotational Program Managers

Define and execute research programs that:

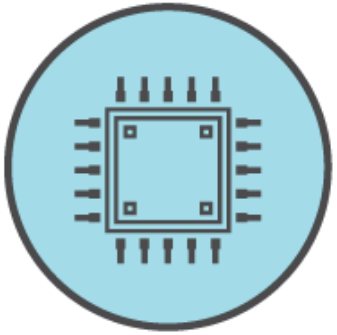
- Have goals that are clear, measureable, ambitious and credible
- Employ independent and rigorous Test & Evaluation
- Involve IC partners from start to finish
- Run from three to five years
- Publish peer-reviewed results and data, to the greatest possible extent
- Transition new capabilities to intelligence community partners

4 Core Research Thrusts

Analysis

Anticipatory Intelligence

Collection



Computing

Analysis R&D

"Maximize insight from the information we collect, in a timely fashion"

LARGE DATA VOLUMES AND VARIETIES

Provide powerful new sources of information from massive, noisy data

SOCIAL, CULTURAL, AND LINGUISTIC FACTORS

Analyze language and speech to produce insights into groups and organizations

IMPROVING ANALYTIC PROCESSES

Enhance analytic process at the individual and group level

Anticipatory Intelligence R&D

"Detect and forecast significant events"

S & T INTELLIGENCE

Detect and forecast the emergence of new technical capabilities

INDICATIONS & WARNINGS

Provide early warning of societal crises, disease outbreaks, insider threats, and cyber attacks

STRATEGIC FORECASTING

Forecast major geopolitical trends and rare events

Collection R&D

"Dramatically improve the value of collected data"

NOVEL ACCESS

Reach hard targets in denied areas

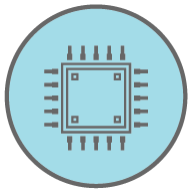
ASSET VALIDATION AND IDENTITY INTELLIGENCE

Assess trustworthiness and advance biometrics in real-world conditions

LOCATING, TRACKING AND DETECTING

Accurately locate and track intelligence interests and detect CBRNE agents

Computing R&D



"Operate effectively in a globally interdependent and networked environment"

COMPUTATIONAL POWER

Revolutionary advances to solve problems intractable with today's computers

TRUSTWORTHY COMPONENTS

Gain the benefits of leading-edge hardware and software without compromising security

SAFE AND SECURE SYSTEMS

Protecting systems against cyber threats

Office of the Director of National Intelligence
IARPA
BE THE FUTURE

How to Engage with IARPA

iarpa.gov | 301-851-7500

info@iarpa.gov

Reach out to our Program Managers.

Schedule a visit if you are in the DC area or invite us to visit you

Opportunities to Engage:

RFIS AND WORKSHOPS

Opportunities to learn what is coming, and to influence programs.

"SEEDLINGS"

Typically a 9-12 month study; you can submit your research proposal at any time. We strongly encourage informal discussion with a PM before proposal submission.

PRIZE CHALLENGES

No proposals required. Submit solutions to our problems – if your solutions are the best, you receive a cash prize and bragging rights.

RESEARCH PROGRAMS

Multi-year research funding opportunities on specific topics.

Molecular Information Storage (MIST)

Program Overview

Office of the Director of National Intelligence

IARPA

BE THE FUTURE

MIST Overview

- MIST is anticipated to be a multi-year research and development program
- The program seeks to develop deployable storage technologies that can eventually scale into the exabyte regime and beyond with reduced physical footprint, power and cost requirements relative to conventional storage technologies.
- MIST seeks to accomplish this by using sequence-controlled polymers as a data storage medium, and by building the necessary devices and information systems to interface with this medium.
- Technologies are sought to optimize the writing and reading of information to/from polymer media at scale, and to support random access of information from polymer media archives at scale

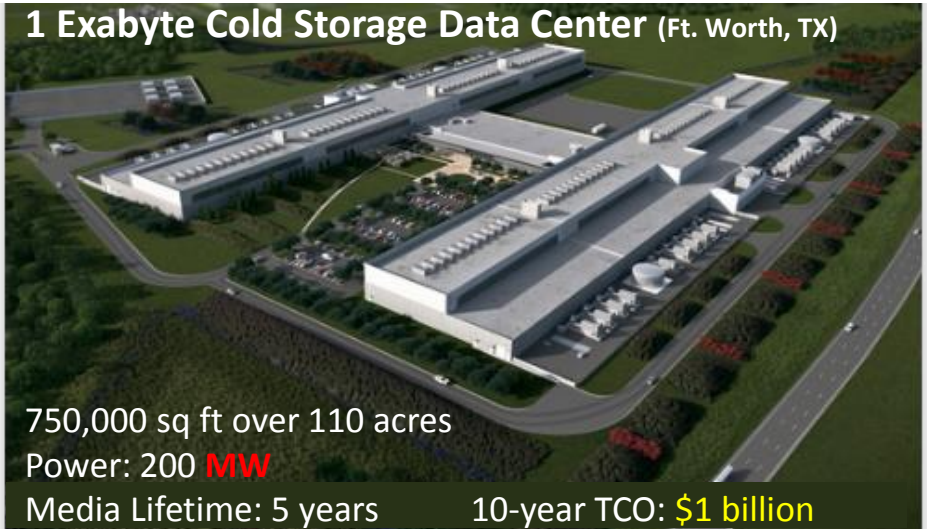
Background

- The scale and complexity of the world's "big data" problems are increasing rapidly.

- Use cases that require storage and random access from exabytes of mostly unstructured data are now well-established in the private sector and are of increasing relevance to the public sector.

- However, meeting these requirements poses extraordinary logistical and financial challenges: today's exabyte-scale data centers occupy large warehouses, consume megawatts of power, and cost billions of dollars to build, operate and maintain over their lifetimes.

Example



TCO = Total Cost of Ownership

- This resource intensive model does not offer a tractable path to scaling beyond the exabyte regime in the future.
- Faced with exponential data growth, large data consumers may soon face a choice between investing exponentially more resources in storage or discarding an exponentially increasing fraction of data.

Background

- Although many factors drive the resource requirements of today's large-scale storage systems, perhaps the single largest factor is media.

- All conventional storage paradigms (magnetic, optical, and solid state) write bit features onto planar media. Once areal storage density has been maximized, these paradigms offer limited ability to write bits isotropically in 3D.

- As a result, to build a data center with exponentially larger capacity than a single unit of planar storage media requires purchasing exponentially more media and read/write hardware, which drives physical footprint, cooling, power and cost requirements.

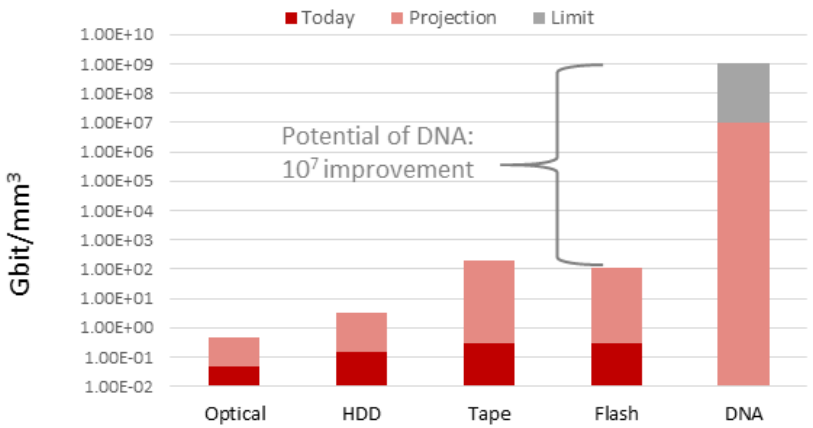
"Wish List" For Next-Gen Storage Media

- 1. Orders of magnitude higher volumetric information density than conventional paradigms, to enable the development of ultra-scalable storage technologies with a substantially smaller footprint, and lower power and cost requirements of associated read/write hardware, than current systems.

- 2. Long-term stability against progressive data degradation, to obviate the need for regular integrity checks and media replacement, and thereby reduce operation and maintenance costs.

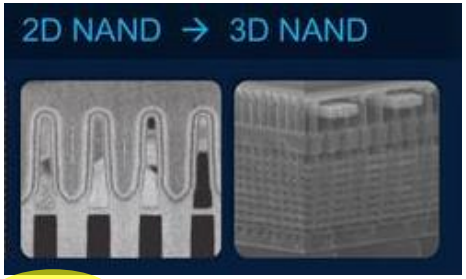
- 3. Basic methods for writing and reading information from the storage medium should already exist, and the engineering optimizations needed to support real-world commercial deployment within a 10 year horizon should be clear and plausible.

Opportunity: Sequence-Controlled Polymers



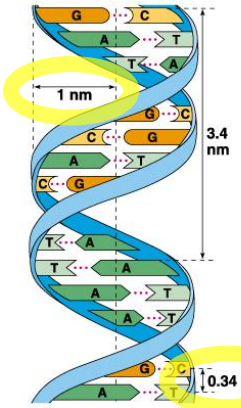
The volumetric information density of conventional storage media vs DNA.

NAND Flash



19 nm features, 64 stacked layers

DNA



- Polymers, or **molecular-scale sequences of physical bits**, can have a stable lifetime of hundreds of years and an information density that is >10^5 times higher than that of conventional storage media (with a raw limit of 1 exabyte/mm^3). This makes polymers attractive as a data storage medium.

Ref: Zhirnov et al, Nature Materials 2016

Data figure courtesy Victor Zhirnov, Semiconductor Research Corporation, and Karin Strauss, Microsoft Corporation

Examples

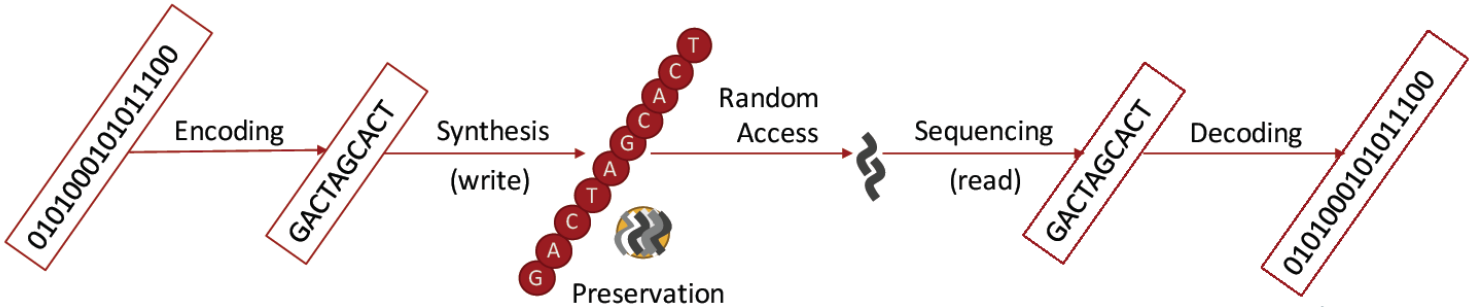


Figure courtesy Luis Ceze, University of Washington

Using DNA as physical media in a molecular information storage system.

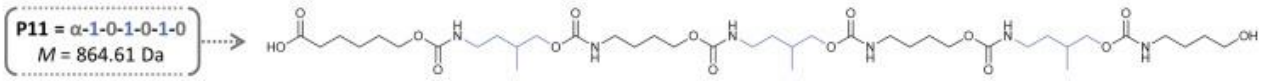
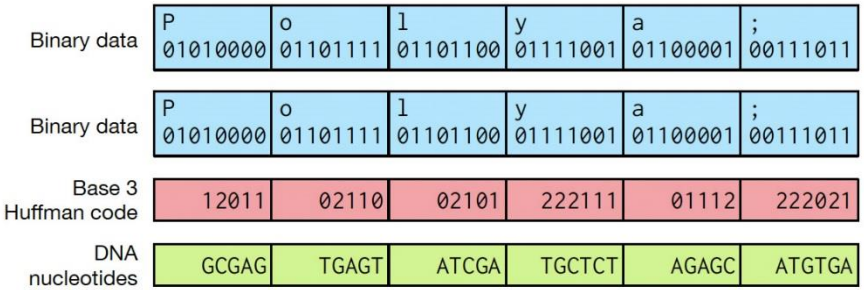


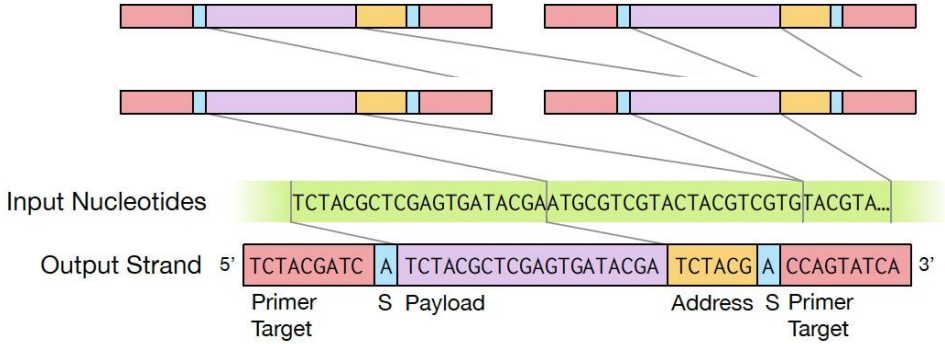
Figure courtesy Jean-Francois Lutz, University of Strasbourg

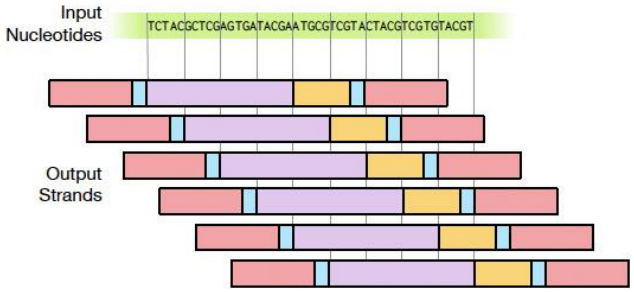
Using commodity polymers to digitally encode sequences.

Encoding Examples

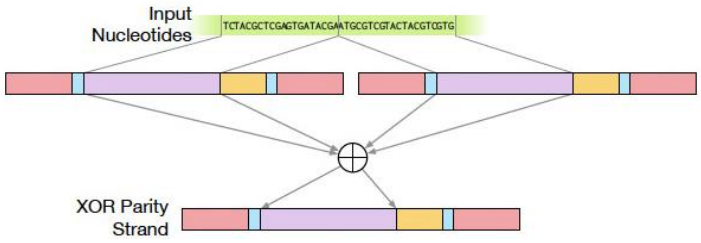


Translating binary data to DNA nucleotides via a Huffman code.





Error correction through 4x redundancy
(Goldman et al, Nature 2013)



Error correction through XOR encoding
(Bornholt et al, ASPLOS 2016)

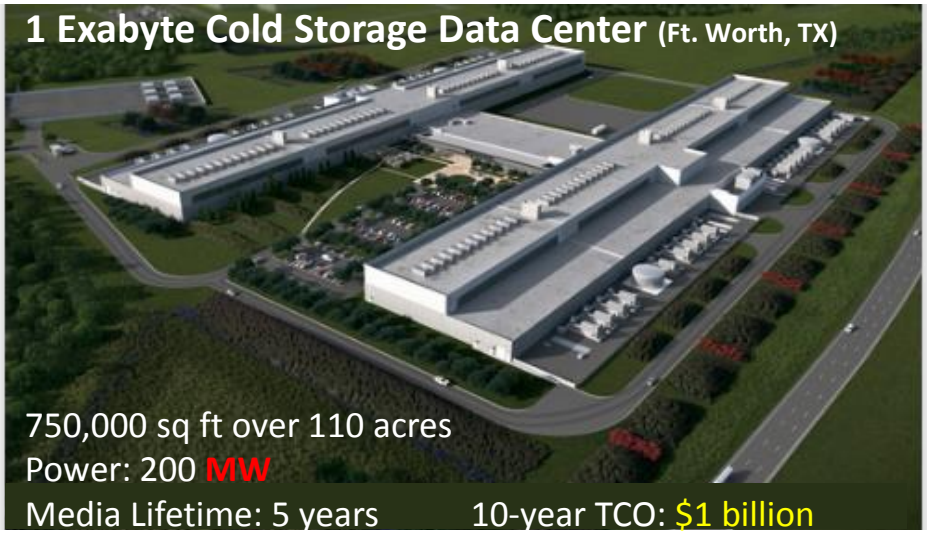
Historical Context

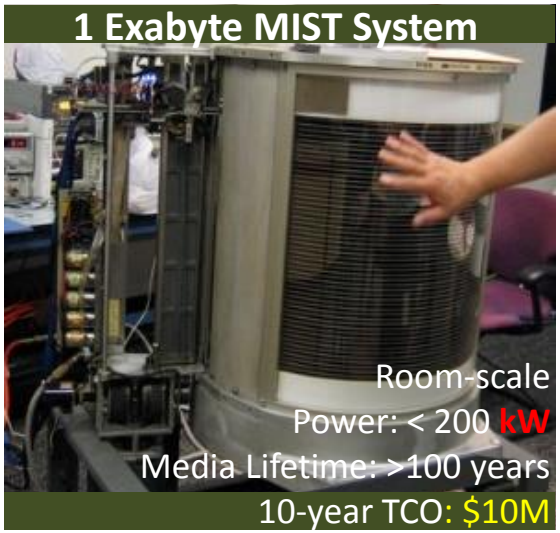
- During 2016 and 2017, IARPA and the Semiconductor Research Corporation organized two workshops that assembled international stakeholders from academia and the biotech, semiconductor and information technology industries to roadmap clear and achievable engineering optimizations that would be necessary to develop scalable MIST systems.
  - Roadmap from 2016 workshop here

- The MIST program now seeks to put this roadmap into practice by assembling a multidisciplinary community around the shared goal of developing compact and scalable molecular information storage technologies to support real-world “big data” use cases.

Program Vision

The end result of the program will be technologies that jointly support end-to-end storage and retrieval at the terabyte scale, and which present a clear and commercially viable path to future deployment at the exabyte scale





TCO = Total Cost of Ownership

Example Approaches

- Innovative solutions are sought and are anticipated to range across a diversity of domains, including chemistry, molecular biology, microfluidics, semiconductor engineering and computer science.

- Example approaches to writing data may include, but are not limited to, performing massively parallel chemical synthesis of polymers on microfabricated chips.

- Example approaches to reading data may include, but are not limited to, sequencing polymers using arrays of nanopore sensors.

- Example approaches to random access may include, but are not limited to, using key-value stores in conjunction with a physical compartmentalization of molecular media by data type.

Current State of MIST Technology

- To date, most work in the MIST space has focused on developing proof-of-concept encoding and decoding schemes for use with DNA storage media.
  - DNA is used for convenience; We understand it well, and biology provides tools for working with it
  - Alternative media, such as peptides or synthetic polymers (e.g. polyimide/maleimide), offer potential advantages (e.g. peptides offer 15 bits/unit vs 2 bits/unit in DNA), but tools are comparatively immature

- Multiple studies have shown DNA can support scalable, random-access and error-free information storage.
  - (Erlich and Zielinski, Science 2017) demonstrated error-correcting codes that approach max capacity of DNA
  - State-of-the-art operating system is a 2016 DNA-based archival storage framework that supports random access from a DNA key-value store (Bornholt at al, ASPLOS 2016)

- Two major categories of technical challenges remain:
  - **Physical Media**: Improving cost, speed, scale of synthesis and sequencing technologies.
  - **Operating System**: Creating scalable indexing, random access and (ideally) search capabilities.

Physical Media Challenges

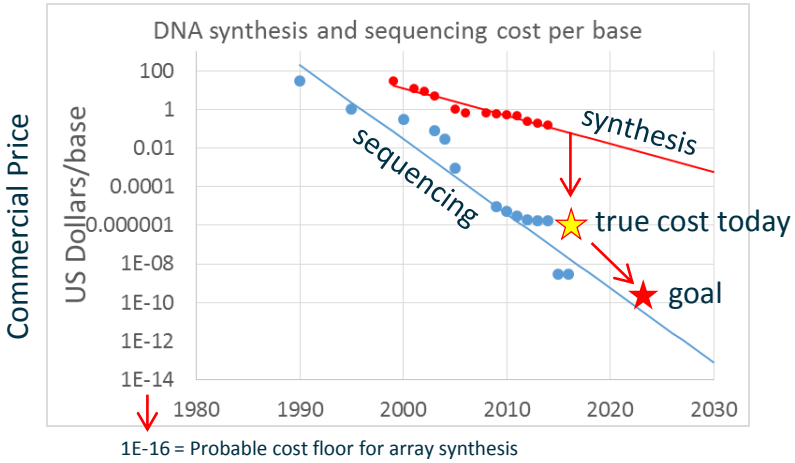
- Improving cost, speed, scale of synthesis and sequencing technologies

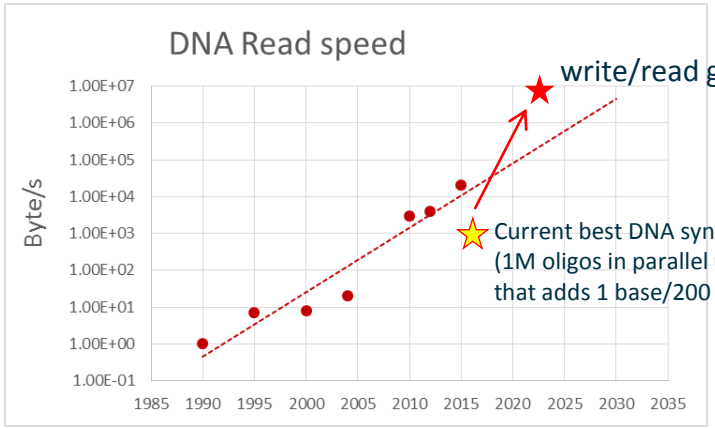
- For DNA, the key challenge is improving these performance measurements beyond the needs of the life sciences industry by several orders of magnitude
  - Life sciences applications require *perfect* synthesis and sequencing; Scale, throughput and cost are secondary design considerations.
  - By contrast, data storage applications can tolerate high write/read error rates; Scale, speed and cost are primary design considerations.

- These challenges are illustrated in the next few slides.

Physical Media Cost, Speed Challenges

- Let's say we want to write 1 TB/day in DNA for $1k, then read it back just as quickly
- Using numbers from 2017 Science paper on DNA Fountain codes:
  - Encoding 2,146,816 bytes required 72,000 oligomers of length 152 nucleotides each
  - That's 10,944,000 bases, or 5.09 bases/byte; Therefore, 1 terabyte should require 5.09 x 10^12 (trillion) bases
  - Writing 1 terabyte/day for $1,000 implies cost/base = 1.9 x 10^-10 $/base, write/read speed = 1.1x10^7 bytes/s





- These are commercial prices, not costs; they include markups for overhead
- Historical data are for life sciences technologies that ensure no errors
- Cost & speed goals are achievable through optimizations.

DNA-Based MIST Workflow Cost, Speed Challenges

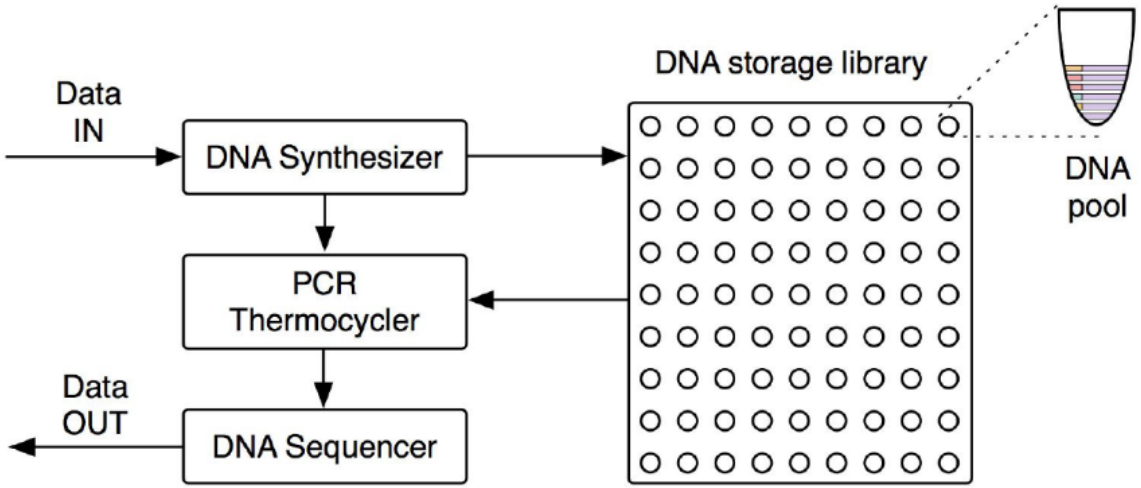
- Currently, too slow and costly to support even *terascale* archival data storage.

Step	Time Required	Cost
1. Encode files to be archived as a set of oligonucleotide sequences in software	Seconds	Negligible
2. Contract with a DNA synthesis company to synthesize many copies of the desired oligos	Multiple weeks from order to receipt of synthesized DNA	1×10^{-4} dollar/base in 2015 (~$3k for 100MB)
3. For each file of interest, identify primers needed to pull it from the archive	Seconds	Negligible
4. Order primers from DNA synthesis company	Days to weeks, depending on the number of primers	A few dollars
5. Hybridize primers to oligos in the data archive; isolate hybridized DNA using gel electrophoresis	Hours	Labor
6. Sequence oligos that were isolated from archive	Hours, if on-site; Days otherwise	1×10^{-9} dollar/base in 2015
7. Decode retrieved files from oligo sequences	Seconds	Negligible

- **Needs:** Solving this problem will require substantial reductions in the cost of DNA synthesis and sequencing, and the deployment of these technologies in a fully automated end-to-end workflow.

DNA-Based MIST Workflow Scale Challenge

- Showing MIST workflow from (Bornholt et al, ASPLOS 2016)



- Storage library included *millimeter* scale vials containing DNA in solution.
- When data from a particular pool were required, this was retrieved manually.
  - Limited scalability: in a 3D library, access to any one vial would require dismantling the entire archive.
  - This is not feasible in a large data regime. Highlights the need for automation and miniaturization.

Some Physical Media Challenges

- To develop practical MIST devices, the cost, speed and scale of synthesis and sequencing technologies must be optimized together in the context of an end-to-end workflow
  - e.g. Require synthesis of X bits & subsequent random access of Y bits from archive of size Z per day
- Challenges related to system integration include:
  - Fluidics automation; reliable interfaces between electronics and wet system components
- Challenges related to parallelization include:
  - Scaling down print-heads; reducing feature sizes to reduce volume for storage/manipulation of media
- Challenges related to read-out include:
  - Non-destructive read-out vs the need to regenerate data after reading

Operating System Challenges

- Creating scalable indexing, random access and (ideally) search capabilities
- How do we index an exabyte archive for fast random access? This is unsolved.
- Given support for random access in physical media, what addressing scheme is optimal? Does the need to optimize random access suggest an optimal physical layout of media?
  - e.g. would it help to physically group molecular media by data type, file size, etc.
- Given support for pattern matching and search in physical media, what encoding is optimal?
  - e.g. if patterns are matched by hybridizing a primer encoding "target features" to stored DNA, how to encode features in DNA to yield maximum versatility and reliability of pattern matching functions?
- Security challenges (e.g. how to manage security policies dynamically?)
- Many of the above will be determined by expected access patterns
  - e.g. archival (reads are uncommon) vs analytics (reads are common)

MIST Program Technical Areas

- **TA1 (Storage):** Develop a table-top device capable of writing information to molecular media with a target throughput and resource utilization budget. Multiple, diverse approaches are anticipated, which may utilize DNA, polypeptides, synthetic polymers, or other sequence-controlled polymer media.

- **TA2 (Retrieval):** Develop a table-top device capable of randomly accessing information from molecular media with a target throughput and resource utilization budget. Multiple, diverse approaches are anticipated, which may utilize nanopores, mass spectrometry, or other methods for sequencing polymers in a high-throughput manner.

- **TA3 (Operating System):** Develop an operating system (OS) for use with storage and retrieval devices that coordinates indexing, addressing, data compression, encoding, error-correction and decoding of files from molecular media in a manner that supports efficient random access at scale. Multiple, diverse approaches are anticipated, which may draw on established methods from the storage industry, or develop new methods to accommodate constraints imposed by polymer media.

Teaming

- Collaborative efforts and teaming among potential performers is highly encouraged.
- It is anticipated that teams will be multidisciplinary, including expertise in fields such as chemistry, molecular biology, microfluidics, semiconductor engineering and computer science.
- Offerors may propose to any combination of one or more TAs, but should plan to be part of an integrated team comprising all three TAs.

Out of Scope

- Approaches that rely on media other than sequence-controlled polymers for long-term data storage are out of scope.

Test and Evaluation

- IARPA will employ a Government Test and Evaluation (T&E) team to assist in evaluating progress and success of the MIST program.

- The T&E team will measure each performer's developed device or OS performance against a set of Metrics and Milestones specific to each technical area.

- Offerors to each TA are strongly encouraged to suggest a test and evaluation methodology that is compatible with the proposed technical approach.

Research Phases

- Phase 1: *Develop storage and retrieval devices and operating system (OS) with performance suitable for gigabyte-scale applications (24 Months)*
  - TA1/TA2 Goals: De-risk scalable synthesis and sequencing approaches
  - TA3 Goals: Develop a simulator of TA1/TA2 hardware; Demonstrate an operating system that supports indexing, addressing and random access at scale in simulation
  - Decision points:
    - Month 12: TA1 delivers a decodable polymer data archive to the government
    - Month 23: TA1/TA2 demonstrate functional devices, TA3 demonstrates operating system capabilities in simulation
  - Output: Fully automated O(10 GB) workflow that can scale after further optimization.

- Phase 2: *Optimize devices and OS to support terabyte-scale applications* (24 Months)
  - Decision points:
    - Month 36: All TAs deliver devices and OS that work together to support O(100GB) workflows
  - Output (Month 48): All TAs deliver devices and OS that work together to support O(1TB) workflows

Test and Evaluation Overview

- MIST Evaluations will occur twice during each phase
- Performers will be evaluated on several metrics
- Technical metrics will become progressively more challenging moving from Phase 1 to Phase 2

Example T&E Metrics & Milestones for TA1

Metric	Milestone
Resource Budget for Storage	Offeror-defined during both phases, but milestone target is a table-top (1 m^2) device with a < $1,000 effective cost of encoding and then writing the following volumes of information to molecular media: ○ Phase 1: 10 GB ○ Phase 2: 1 TB Resource budget may entail cost of reagents, volume and/or cost of safe disposal of waste products, power requirements, or other factors. Resource budget may be extrapolated from a demonstration involving less data than specified above.
Write Throughput	Using an encoding/decoding scheme of the performer's choice, achieve the following write throughput: ○ Phase 1: 10 GB/day ○ Phase 2: 1 TB/day Write throughput may be extrapolated from a demonstration involving less data in less time.
Device Storage Capacity	○ Phase 1: 10 GB ○ Phase 2: 1 TB
Maximum Volumetric Information Density	Offeror-defined in both phases
Write Error Rate	Offeror-defined in both phases.

Likely T&E Methodology for TA1

- To assess performance, T&E partners may require physical access to devices and the ability to instrument them with sensors.

- To assess synthesis performance, T&E partners will provide TA1 performers with a digital collection of files that must be written to polymer media within 24 hours, after which T&E partners will take possession of the polymer media archive for sequencing and decoding using methods identified in consultation with performers.

  - In Phase 1, the digital collection of files to be stored will include a mixture of both structured and unstructured data, including text documents, spreadsheets, server logs, images, audio and video, with file sizes ranging from kilobytes to megabytes.
  - In Phase 2, the digital collection of files to be stored will include the same diversity of file types and sizes used in Phase 1, but will further be tailored for specific use cases that have relevance to program transition partners. Example use cases include storage and random access retrieval from archives containing audio and video, transactions, genomes, and/or neuroscience data.

Example T&E Metrics & Milestones for TA2

Metric	Milestone
Resource Budget for Retrieval	Offeror-defined during both phases, but milestone target is table-top (1 m^2) device with a <$1,000 cost of retrieving and decoding the following volumes of information from molecular media: ○ Phase 1: 1 TB ○ Phase 2: 10 TB O(TB) to lay groundwork for analytics applications Resource budget may entail cost of reagents, volume and/or cost of safe disposal of waste products, power requirements, or other factors. Resource budget may be extrapolated from a demonstration involving less data than specified above.
Read Throughput	Using an encoding/decoding scheme of the performer's choice, achieve the following read throughput: ○ Phase 1: 1 TB/day ○ Phase 2: 10 TB/day O(TB) to lay groundwork for analytics applications Read throughput may be extrapolated from a demonstration involving less data in less time.
Read Error Rate	Offeror-defined in both phases.

Likely T&E Methodology for TA2

- To assess performance, T&E partners may require physical access to devices and the ability to instrument them with sensors.

- To assess random access and decoding accuracy, T&E partners will provide TA2 performers with a polymer data archive that must be sequenced and decoded within 24 hours, after which T&E partners will evaluate the accuracy of decoded data with reference to source data that were used to generate the molecular archive.

- To enable appropriate planning by T&E partners, offerors must specify their requirements concerning polymer chemical composition, addressing and encoding scheme, copy number, physical organization, or other properties of the source molecular archive that will be provided to TA2 for read-out.

- In both phases, the composition of the digital collection of files to be recovered will be similar to the collection specified for TA1 above.

Example T&E Metrics & Milestones for TA3

Metric	Milestone
Resource Requirements and Performance Characteristics of Simulated Storage and Retrieval Hardware	Offeror-defined during both phases. It is anticipated that this metric will be refined in consultation with T&E partners who will use the simulator to evaluate performer operating systems.
Resource Requirements for Operating System-Specific Steps of End-to-End Storage and Retrieval Workflow	Offeror-defined during both phases, but required hardware must fit within a standard 2U form factor. Requirements must be specified separately for encoding and decoding steps of the workflow.
Precision and Recall of Random Access Operations	Offeror-defined during both phases. It is anticipated that this metric will be refined in consultation with T&E partners who will use the simulator to evaluate performer operating systems.
Other Metrics	Offeror-defined. Examples may include bit error rate, interface parallelism, write/read bandwidth, latency, etc.

Likely T&E Methodology for TA3

- In Phase 1, T&E partners will evaluate operating systems that are interfaced with simulated write and read hardware provided by performers. T&E will explore the robustness of indexing, random access and encoding/decoding operations to probable hardware failure modes in simulation.

- In Phase 2, T&E partners will further evaluate operating systems that are interfaced with hardware simulators, and by the end of the phase, will also evaluate operating systems that are interfaced with physical storage and retrieval hardware.

- In both phases, the composition of the digital collection of files to be stored and retrieved by TA3 will be similar to the collections specified for TA1 and TA2 above. However, the composition and size of the collection may be changed to support evaluation of scalability.

Summary

- We are looking for a diversity of approaches to developing deployable storage technologies that can eventually scale into the exabyte regime and beyond with reduced physical footprint, power and cost requirements relative to conventional storage technologies.
- We anticipate teams will include individuals with expertise/experience in chemistry, molecular biology, microfluidics, semiconductor engineering and computer science.
- We expect teams will have a strong plan for working across team members to accomplish the program goals
- Please do review this brief, the draft BAA and any other materials on the MIST website. Please send us your recommendations and suggestions for the BAA – using the prescribed format.
- The BAA will supersede anything presented or said at this Proposers' Day by IARPA.

Point of Contact

Dr. David A. Markowitz

Program Manager

IARPA, Office of the Director of National Intelligence

Intelligence Advanced Research Projects Activity

Washington, DC 20511

Phone: (301) 851-7497

Fax: (301) 851-7673

Electronic mail: dni-iarpa-baa-18-03@iarpa.gov

(include IARPA-BAA-18-03 in the Subject Line)

Website: www.iarpa.gov

Questions? Please fill out cards.

Doing Business with IARPA

Acquisition Team

Office of the Director of National Intelligence

IARPA

BE THE FUTURE

Responding to Q&As

- Please read entire BAA before submitting questions
- Pay attention to Section 4 (Application & Submission Info)
- Read Frequently Asked Questions on the IARPA @

  http://www.iarpa.gov/index.php/faqs

- Send your questions as soon as possible
  - 18-03 BAA: **dni-iarpa-baa-18-03@iarpa.gov**
  - Write questions as clearly as possible
  - Do NOT include proprietary information

Eligible Applicants

- Collaborative efforts/teaming strongly encouraged
  - Content, communications, networking, and team formation are the responsibility of Proposers
- Foreign organizations and/or individuals may participate
  - Must comply with Non-Disclosure Agreements, Security Regulations, Export Control Laws, etc., as appropriate, as identified in the BAA

Ineligible Organizations

Other Government Agencies, Federally Funded Research and Development Centers (FFRDCs), University Affiliated Research Centers (UARCs), and any organizations that have a special relationship with the Government, including access to privileged and/or proprietary information, or access to Government equipment or real property, are not eligible to submit proposals under this BAA or participate as team members under proposals submitted by eligible entities.

Intellectual Property (IP)

- Unless otherwise requested, Government rights for data first produced under IARPA contracts will be UNLIMITED
- At a minimum, IARPA requires Government Purpose Rights (GPR) for data developed with mixed funding
- Exceptions to GPR
  - State in the proposal any restrictions on deliverables relating to existing materials (data, software, tools, etc.)
- If selected for negotiations, you must provide the terms relating to any restricted data or software, to the Contracting Officer

Pre-Publication Review

- Funded Applied Research efforts, IARPA encourages:
  - Publication for Peer Review of **UNCLASSIFIED** research

- Prior to public release of any work submitted for publication, the Performer will:
  - Provide copies to the IARPA PM and Contracting Officer Representative (COR/COTR)
  - Ensure shared understanding of applied research implications between IARPA and Performers
  - IARPA PM decides on approval for release or receiving courtesy copy

Preparing the Proposal

- Note restrictions in BAA Section 4 on proposal submissions
  - Interested Offerors must register electronically IAW instructions on: **https://iarpa-ideas.gov**
  - Interested Offerors are strongly encouraged to register in IDEAS at least 1 week prior to proposal "Due Date"
  - Offerors must ensure the version submitted to IDEAS is the "Final Version"
  - Classified proposals – Contact IARPA Chief of Security
- BAA format is established to answer most questions
- Check FBO for amendments & IARPA website for Q&As
- BAA Section 5 – Read Evaluation Criteria carefully
  - e.g. "The technical approach is credible and includes a clear assessment of primary risks and a means to address them"

Preparing the Proposal (BAA Sect 4)

- Read IARPA's Organizational Conflict of Interest (OCI) policy: **http://www.iarpa.gov/index.php/working-with-iarpa/iarpas-approach-to-oci**
- See also eligibility restrictions on use of Federally Funded Research and Development Centers, University Affiliated Research Centers, and other similar organizations that have a special relationship with the Government
  - Focus on possible OCIs of your institution as well as the personnel on your team
  - See Section 4: It specifies the non-Government (e.g., SETA, FFRDC, UARC, etc.) support we will be using. If you have a potential or *perceived* conflict, request a waiver as soon as possible

Organizational Conflict of Interest (OCI)

- If a prospective offeror, or any of its proposed subcontractor teammates, believes that a potential conflict of interest exists or may exist (whether organizational or otherwise), the offeror should promptly raise the issue with IARPA and submit a waiver request by e-mail to the mailbox address for this BAA at **dni-iarpa-baa-XX-XX@iarpa.gov**.
- A potential conflict of interest includes but is not limited to any instance where an offeror, or any of its proposed subcontractor teammates, is providing either scientific, engineering and technical assistance (SETA) or technical consultation to IARPA. In all cases, the offeror shall identify the contract under which the SETA or consultant support is being provided.
- Without a waiver from the IARPA Director, neither an offeror, nor its proposed subcontractor teammates, can simultaneously provide SETA support or technical consultation to IARPA and compete or perform as a Performer under this solicitation.

Streamlining the Award Process

- Cost Proposal – we only need what we ask for in BAA
- Approved accounting system needed for Cost Reimbursable contracts
  - Must be able to accumulate costs on job-order basis
  - DCAA (or cognizant auditor) must approve system
  - See **http://www.dcaa.mil** , "Audit Process Overview - Information for Contractors" under the "Guidance" tab
- Statements of Work (format) may need to be revised
- Key Personnel
  - Expectations of time, note the Evaluation Criteria requiring relevant experience and expertise
- Following selection, Contracting Officer may request your review of subcontractor proposals

IARPA Funding

- IARPA funds Applied Research for the Intelligence Community (IC)
  - IARPA cannot waive the requirements of Export Administrative Regulation (EAR) or International Traffic in Arms Regulation (ITAR)
  - Not subject to DoD funding restrictions for R&D related to overhead rates
- IARPA is not DoD

Disclaimer

- Content of the Final BAA will be specific to this program
  - The Final BAA is being developed
  - Following issuance, look for Amendments and Q&As
  - There will likely be changes
- The information conveyed in this brief and discussion is for planning purposes and is subject to change prior to the release of the Final BAA.

Point of Contact

Dr. David A. Markowitz

Program Manager

IARPA, Office of the Director of National Intelligence

Intelligence Advanced Research Projects Activity

Washington, DC 20511

Phone: (301) 851-7497

Fax: (301) 851-7673

Electronic mail: dni-iarpa-baa-18-03@iarpa.gov

(include IARPA-BAA-18-03 in the Subject Line)

Website: www.iarpa.gov

Questions? Please fill out cards.