

**Deep Exploration and Filtering of Text (DEFT)
Questions and Answers
As of May 30, 2012**

Note: the below Q/A are from the Proposers' Day that was held on May 16, 2012 and describe DARPA's current plan for the DEFT program technical areas (TAs) as referenced in the Proposers' Day presentations (posted as a separate document). The answers to some of the questions have not yet been determined and are shown as "TBD." As a result of the publication of DARPA-BAA-12-47 on May 22, 2012, answers to some questions have been updated from the initial Proposers' Day responses.

	Question	Answer
103	Are evaluations expected to involve human use?	No. Human subject research is not anticipated for this program. However, if it is proposed, proposers must follow the requirements as described in the BAA. Proposers should also be aware that it can take up to 6 months to complete the approval process for efforts involving human use.
102	Does the system require an open source technology solution, or is proprietary okay?	This is up to the proposer. If an approach is based heavily on proprietary information, you will need to explain how we can still meet our goals of serving our transition partners. See specifics in the BAA.
101	Will there be a program-wide knowledge architecture, formal ontology or other representation to cover the examples from your briefing slides, and will it be developed incrementally during the DEFT program?	The TA2 performer should suggest a knowledge architecture and other formal representations as specified in the BAA.
100	Is it the case that TA1 is focused entirely on HLT algorithms, or does it also include other types of algorithms that span additional areas (such as joint inference)?	HLT algorithms relevant to DEFT goals are central to TA1, as well as algorithms for joint inference and Alert Generation.
99	Does the TA2 performer need to provide graphical or data visualization?	TA2 proposals should suggest basic GUI's as specified in the BAA.

	Question	Answer
98	Are entity relations in scope? Is there an interest in group detection and perspective?	Yes. If what is meant is opinion vs. fact, then yes.
97	Should the system automatically learn what is of interest to the user?	Proposals should rely on linguistic content only and not be dependent on any particular type of user input.
96	Will the end user specify what is of interest to him/her as input to the system?	TBD.
95	Is the focus on novel algorithms or upon novel implementations of existing algorithms?	Novel algorithms.
94	Assuming there will be multiple report inputs, do source/origin/author data matter? Will they be available?	The amount and type of meta data will be determined by DARPA during the program. Proposals should rely on linguistic content only and not be dependent on any particular type of meta-data.
93	Should we assume report and speech text input are authentic (i.e., not deceptive)?	No.
92	Is a subcontractor plan required if the value is greater than \$650k and it will be awarded under a cooperative agreement?	Not for a cooperative agreement. Subcontractor plans are only required for FAR-based procurement contracts.
91	Will access to the internet be permitted for the DEFT prototype?	TBD.
90	What about the use of reference databases (e.g., dbpedia, Wikipedia, etc.)?	You can use outside data resources, but TA1/TA2 performers will not be funded to create/collect data, and there is risk of mismatch with respect to data in a classified setting.
89	Linking information requires reasoning either at a language level or at the domain level or both. Would the data be domain-specific?	Domain-specific, but not task-specific--DoD or Intelligence Community (IC) focused.
88	Will the data conform to a limited set of orthographies (writing style/format)?	TBD.
87	Does DARPA have a favored machine translation tool?	No.

	Question	Answer
86	Are approaches involving semi-supervised feedback from analysts in scope?	Yes.
85	How will DARPA facilitate relationships with transition partners? Or is TA2 responsible for developing those relationships?	DARPA will identify transition partners at program kick-off.
84	Can a single team propose several TA1 algorithms in one proposal? And then also to TA3?	Yes, up to five algorithms may be proposed for TA1. You may propose to both TA1 and TA3, but you will not be selected for both TA1 and TA3. Specifics are in the BAA.
83	Is the automated pre-report/summary creation task only included in TA2? Can TA1 proposers work on this as well?	Implemented algorithms in TA1 may be proposed to support pre-report/summary creation.
82	Will the domain remain in the general IC area?	It is DoD and IC.
81	Will this be a 6.2 or 6.3 program?	DARPA anticipates 6.2 funding.
80	What format will data be provided in?	TBD.
79	Will data be provided for all TAs or is data acquisition required?	Data will be provided to TA1 and TA2.
78	What deliverables do you anticipate from TA3? Is that in terms of words? Or ground truth?	See the TA3 description and the Deliverables sections in the BAA for specifics.
77	Could you clarify the kinds of entities and relations that will be focused on through the program? Is it unlimited or will it be focused?	TBD but generality is preferred.
76	Does DEFT have anything to do with recent AUKS workshop?	Many HLT-related workshops and conferences are relevant to DEFT.
75	Should multiple teams develop or use a common KR&R framework? How would this be coordinated?	A common KR&R framework is not required.
74	Do TA1 performers need to be allied with the TA2 performer?	No. See Q/A 15.

	Question	Answer
73	In language, “everything depends on everything else”. Do you imagine TA1 proposals that simultaneously address multiple HLT algorithms in an integrated way (example: semantic engineering and coreference and relational analysis) as opposed to leaving that up to TA2?	We leave it open as to how TA1 proposals will present algorithms and relationships between algorithms.
72	How important will it be that algorithms can scale? Should scalability be built-in from inception or is that secondary?	Yes, scalability is important.
71	Is there an architectural baseline we should keep in mind when building analytics? For instance, must we be able to interoperate with cloud-based resources?	There are no requirements at this time.
70	Can TA1 proposers propose options?	Yes.
69	Will there be any throughput requirements for TA1 performers (e.g., characters analyzed per second)?	Throughput requirements will not be a part of proposal evaluation. The only requirement of TA1 implemented algorithms is that they run with the TA2 platform according to the TA2 API.
68	Would there be interest in modeling the emotional tone and dynamics of conversation?	Yes.
67	Could you please repeat the annotation categories for the data in TA3?	See the TA3 description and the Deliverables sections in the BAA for specifics.
66	For TA1 algorithms are there any assumptions being made or constraints/requirements regarding the data sources over which the algorithms will be executed (e.g., Hadoop, RDB, etc.)?	No.
65	Will the TA2-built DEFT system have a structured data component and/or an interface to structured data?	Yes.

	Question	Answer
64	Can TA1 algorithms use accumulated prior knowledge as a systematic component of their performance?	Yes.
63	Is there an anticipated Phase 1 down select? If so, what is the weighting of evaluations in that decision?	No. See Q/A 28.
62	Will you be requesting abstracts ahead of proposal submissions?	No.
61	Can you restate the number of performers for each TA?	DARPA anticipates multiple performers for TA1, one for TA2, and one or more for TA3.
60	Is the capability to reside in theater sites? What are the environments anticipated for end-user sites?	TBD.
59	Is the TA2 performer responsible for certification and accreditation on the end-user/transition partner network?	Yes.
58	Is there an anticipated need for multi-level secure corpus and/or knowledge base?	Multi-level security is not a part of this program.
57	Is the enrichment of the corpus and/or knowledge base purely to be machine derived or are human annotations to be incorporated?	Enrichment of the corpus or knowledge base should be automatic.
56	If a team proposes to TA1 with less than 5 algorithms focused on a specific area (such as Relational Analysis), will the algorithms be evaluated separately/independently or might they be evaluated jointly?	If you mean how your proposal will be evaluated, we will evaluate all algorithms submitted; we reserve the right to select all, some, or none.

	Question	Answer
55	It looks like algorithm outputs will be consumed directly by humans/analysts. Are you considering any more structured result representations (e.g., databases, RDF, etc.)?	Yes.
54	Is there a target award date?	For budgeting purposes, proposers may assume an estimated start date of October 1, 2012.
53	If a team has an approach that spans HLT elements (e.g., Filtering and Relation Analysis) will the evaluation support “joint evaluation” across elements?	TBD by NIST and DARPA.
52	Are standards for logical representation set by NIST or the TA2 performer?	NIST will not define this. TA2 will set the standards in coordination with the TA1 performers and DARPA.
51	Does the end-to-end DEFT system have to integrate directly with transition partners’ software/applications/systems? Or is the DEFT system a standalone system?	Both --- NIST closed evaluation expects a standalone end-to-end system; at the end-users’ sites the expectation is integration with end-users’ software/applications/systems.
50	Are neurotechnology solutions using the human as the sensor prior to the alert output of interest (e.g., DARPA’s Neurotechnology for Intelligence Analysts (NIA) program)?	No.
49	Can proposers propose algorithms without an NLP system?	Yes, but the algorithms must be implemented.
48	Can proposers propose their current NLP system for DEFT?	NLP advances can be leveraged, but proposed research should investigate innovative approaches that enable revolutionary research. Specifically excluded are evolutionary advances to the existing state of practice.
47	Can NIST TAC participants propose to DEFT?	Yes, participation in NIST TAC does not create an OCI restriction. Anyone may propose as long as no eligibility restrictions exist.

	Question	Answer
46	How does DEFT relate to current NIST programs such as TAC?	TAC and/or TREC will be the community in which we will situate these evaluations.
45	Is there any seedling money attached to DEFT? If so, how much?	No.
44	How will granularity be addressed? X could cause Y generally, inevitably, or via some other mechanism Z. How should such relations be reported?	This decision is up to the proposer.
43	How many foreign languages will there be?	One or more.
42	How will pre-report summaries be evaluated?	Community-standard summarization measures will be used.
41	Are TA3 proposals considered to be research proposals or a component of testing and evaluation? For example, could an FFRDC or UARC that is prohibited from submitting research proposals to DARPA submit a TA3 proposal?	TA3 proposals are data providers that support the overall research effort. DARPA does not prohibit eligible FFRDCs and UARCs from submitting proposals against this BAA. It is the FFRDC/UARC's responsibility to establish their eligibility to propose to a Government solicitation.
40	In TA1, can proposals include advanced hardware technology that accelerates a proposed algorithm?	It can be proposed, however, the delivery to TA2 is software only.
39	When will the BAA be published?	DARPA-BAA-12-47 was published to FedBizOpps and Grants.gov on May 22, 2012.
38	Is it correct that a single TA1 performer may propose as many as 5 algorithms (but do not need to propose more than 1)?	Yes.
38	What is the expected team size for TA1 performers? Do you expect multi-institution teams?	DARPA has no expectation or preference.
36	If something is proposed to TA1 that you didn't think of, how will data annotation happen?	TA1 proposers may suggest data annotation requirements that support their approach.
35	Are TA1 performers evaluated in terms of how much they contribute to TA2? If so, how should "contribution" be understood?	TA1 implemented algorithms will be evaluated in each phase in an Open NIST Evaluation and are also evaluated with respect to their impact in the integrated end-to-end DEFT System developed by the TA2 performer.

	Question	Answer
34	To what extent is DEFT interested in algorithms or alert types specifically designed to support better analyst behaviors (e.g., reduction of cognitive bias, integration of sentiment)?	Analyst behaviors are not a priority for DEFT.
33	Are TA2 performers supposed to have relationships with possible transition partners? Will those relationships determine the transition partners?	No. Transition partners will be identified by DARPA at program kick-off.
32	What are the target scores at the end user sites?	F-score will be used to determine correctness of Alerts at the cluster level, where precision refers to how many of the found clusters are correct and recall refers to how many correct clusters in the full space were found. Each end user will apply a site-specific threshold on precision and recall. (It should be expected that certain sites will weigh precision more heavily than recall, whereas others will prefer the opposite weighting.) In addition, operational utility of the Alerts, event-focused annotations, and the pre-report summary bullets will be validated by the end user.
31	Who is the evaluator at the end user sites?	An intermediary (an FFRDC or other such organization) identified by DARPA will assist the end user with evaluation. In addition, the FFRDC will collaborate with the TA2 performer to help interpret evaluation results for TA1 performers who cannot work in classified environments.
30	Is there a way to bid on evaluation?	No.
29	Are TA1 or TA2 performers delivering a piece of hardware to an evaluator?	There are no plans for hardware deliveries at this time.
28	What competition will there be? Will there be Go/No-Go's or down-selects?	No, there are no Go/No-Go's or phase metrics and DARPA does not anticipate any down-selects.
27	Can TA1 and TA2 performers translate the foreign language data themselves if they so desire?	Yes, for development.
26	How will you decide the foreign language(s)?	DARPA will select the language(s) after discussion with transition partners and performers.

	Question	Answer
25	What access to classified data will be required/allowed?	All prime proposers to Technical Area 2 must have a Top Secret facility clearance and have personnel under their CAGE code with a Top Secret clearance that are eligible for SCI. It is not necessary for all members of a prime proposer's team (including subcontractors) to hold clearances; however, all primes are responsible for ensuring that un-cleared personnel are properly firewalled from access to classified data. Specific details will be provided in the BAA. Proposers for Technical Areas 1 and 3 are not required to have security clearances.
24	Can TA1 and TA2 performers use their own data?	Yes, for development.
23	How can TA2 integrate such a wide variety of TA1 algorithms that run on different platforms?	In all three phases, algorithms will be developed and implemented on the basis of application programming interfaces (APIs) provided to TA1 performers by the TA2 performer. Algorithm development and implementation is expected to be iterative and cooperative between TA1 and TA2, with three deliveries of implemented algorithms by the TA1 performers to the TA2 performer in each phase. Those algorithms that perform well in NIST open evaluations will be integrated (by the TA2 performer) into the end-to-end DEFT system and are expected to contribute to overall performance improvements in both NIST closed evaluations and transition evaluations.
22	How comprehensive should proposed implemented algorithms be?	A single, implemented algorithm may encapsulate as much as the full capability of an HLT component or it may consist of a sub-capability that is evaluable within a paradigm to be proposed according to the NIST Open Evaluation requirements. In addition, implemented algorithms may be proposed that are not tied to a specific HLT component but contribute to the overall DEFT System (e.g., Alert Generation).
21	Are TA2 performers supposed to deliver hardware, software, or both to transition partners?	Software.
20	Are TA1 performers supposed to deliver hardware, software, or both to the TA2 performer?	Software.

	Question	Answer
19	Are TA1 performers supposed to deliver hardware, software, or both to NIST?	Software.
18	Will there be a delayed start of other program activities for data collection to take place?	No. Note, data is being created as opposed to collected.
17	Should primes be companies or can they be other entities?	DARPA has no preference with regard to type of entity (e.g., industry, university, etc.).
16	Do you expect large teams with big company primes or individual organizations as primes?	For TA1, multiple individual teams (of varying sizes) are anticipated---the prime may be a university, a small company, or a large company and there may or may not be subcontracts. For TA2, a single prime is anticipated; any qualified proposer (regardless of entity type) could be that prime and there may or may not be subcontracts; For TA3 there will be one or more data providers and there may or may not be subcontracts.
15	Do TA1 proposers need to team with (or subcontract to) a TA2 integrator?	No. While TA1 and TA2 performers will be expected to work together, they are separate technical areas and should be proposed separately.
14	How many TA2 performers will there be?	One is anticipated
13	What is the difference between the initial closing and the final closing?	The BAA will remain open until the final closing date/BAA expiration, but proposers are warned that the likelihood of funding is greatly reduced for proposals submitted after the initial closing date. See specifics in the BAA.
12	Should subcontractors submit a detailed technical proposal to you?	For the Technical proposal (i.e., Statement of Work), the answer is No. The Prime should submit a proposal which incorporates all work to be performed by the team and complies with the page limit rules. For the Cost volume, the prime should include sanitized subcontractor costs. Subcontractors who wish to submit a proprietary cost proposal should send it to the BAA mailbox (DEFT@darpa.mil).
11	Who are the transition partners?	Transition partner(s) will be announced at program kick-off.
10	Is international participation allowed?	Yes. Non-U.S. organizations and/or individuals may participate to the extent that such participants comply with any necessary nondisclosure agreements, security regulations, export control laws, and other governing statutes applicable under the circumstances.

	Question	Answer
9	Who is the agent?	The Air Force Research Laboratory.
8	If we are submitting to more than one technical area, do we submit more than one proposal?	Yes. Do not combine different technical areas into one proposal---a separate proposal should be submitted for each technical area. Please heed the page limits outlined in the BAA. Note that TA2/3 proposers will have 15 pages for the Technical volume, and TA1 proposers will have 15 pages if they are proposing one implemented algorithm (and one additional page per algorithm, for no more than a total of five implemented algorithms).
7	How much money is available to TA3?	This will be determined by the quality of proposals.
6	How much money is available to TA2?	This will be determined by the quality of proposals.
5	How much money is available to TA1?	This will be determined by the quality of proposals.
4	How much money is there in the program?	No comment.
3	How many TA1 performers will there be?	Multiple TA1 performers are anticipated.
2	If we are bidding on TA3, can we bid on TA1 or TA2?	Yes, you can propose on all three but can't win all three. The BAA has specific information regarding eligibility.
1	Can a TA2 performer also be a TA1 performer?	Yes however, separate TA1 and TA2 proposals are required.