



Request For Information – DNAtoFace

RFI Number: IARPA-RFI-17-01

Agency: Office of the Director of National Intelligence

Office: Intelligence Advanced Research Projects Activity

IARPA-RFI-17-01

Synopsis

Request for Information (RFI): DNAtoFace

The Intelligence Advanced Research Projects Activity (IARPA) is seeking information on using genetic information to predict facial structure phenotype. This request for information (RFI) is issued solely for information gathering and planning purposes; this RFI does not constitute a formal solicitation for proposals. The following sections of this announcement contain details of the scope of technical efforts of interest, along with instructions for the submission of responses.

Background & Scope

DNA is used by multiple components of the United States (U.S.) government for identifying individuals, detecting and diagnosing medical conditions, and for determining family relationships. Depending on the circumstances, DNA samples can be obtained through consent from a witting individual, or they can be forensically recovered from left behind cells. All cells (e.g., skin, hair, blood, saliva, etc.) contain DNA. In support of scenarios where a forensic DNA sample is recovered from an unknown person, IARPA would like to investigate using genetic DNA phenotyping to provide investigative or intelligence leads for identifying the person by providing a possible face structure and appearance of the unknown person.

Advancements in genetic phenotyping suggests the possibility of predicting a human's facial structure or other attributes from DNA sequences. IARPA is interested in knowing whether single nucleotide polymorphisms (SNP) yield sufficient information for making such prediction or if the whole genome sequence is required. If it is possible to develop a robust genotype to phenotype face prediction model, additional information will be required to determine the required sample size of the training dataset. The DNA training dataset would be used for developing the technology or algorithms required to create the training models. An additional ground truth dataset that contains known face images of the individual will be required to then test the accuracy of the training models and predictive algorithms. These DNA SNP and whole genotype databases, along with the corresponding ground truth face images, would need to be created if not already available for government use. Consideration must be made for determining the appropriate population that should be solicited for developing the training and testing datasets.

Responses to this RFI should answer any or all of the following questions:

1. What is the maturity and level of accuracy of genetic phenotyping outside of gender and genetic ancestry? Is additional information required to phenotype specific characteristics (e.g., height, eye color, skin tone, face structure, etc.)?

2. Who are the major government, industry, and academic leaders in the field of genetic phenotyping?
3. Compare and contrast the leading approaches and techniques for genetic phenotyping. Are any commercial capabilities available?
4. What level of confidence are geneticists, scientists, or researchers able to predict major phenotype information (e.g., height, eye color, skin tone, face structure, etc.) from a whole DNA sequence? Is additional information outside the whole DNA sequence required?
5. What is the impact of utilizing SNP as opposed to whole genome sequencing for predicting genotype to phenotype? Specifically, can facial structure prediction (phenotyping) be achieved with using just SNP? Will (and how will) this limit the accuracy of the predictions? Does the number of SNP collected (e.g., 500,000, 1,000,000, or 5,000,000) impact ability to predict a phenotype from genotype? Which specific SNP should be captured for a face structure phenotype prediction?
6. How many subjects are needed to train a model to predict facial structure and appearance from both SNP and whole genome sequences? Is the required sample size different for SNP versus whole genome sequencing? Does ethnicity, age, or gender impact the required number of subjects?
7. What large-scale SNP or whole genome sequence databases are available in government, academia, and industry? Do they contain corresponding face images? What are the terms of use for such databases?
8. How will epigenetic factors play into any resulting analysis of attributes? What types of epigenetic tests and methodology should be considered?
9. What types of statistical analysis have been done utilizing methods such as power analysis to determine how many subjects are needed to analyze non-disease based phenotypes? Please identify any research (peer reviewed or otherwise) that addresses the sampling needs from a theoretical or quantitative perspective.
10. What other issues do you feel are important to being able to predict non-disease phenotypes from genetic information?

Preparation Instructions to Respondents

IARPA requests that respondents submit ideas related to this topic for use by the Government in formulating a potential program. IARPA requests that submittals briefly and clearly describe the potential approach or concept, outline critical technical issues/obstacles, describe how the approach may address those issues/obstacles and comment on the expected performance and robustness of the proposed approach. If appropriate, respondents may also choose to provide a non-proprietary rough order of magnitude (ROM) regarding what such approaches might require in terms of funding and other resources

for one or more years. This announcement contains all of the information required to submit a response. No additional forms, kits, or other materials are needed.

IARPA appreciates responses from all capable and qualified sources from within and outside of the U.S. Because IARPA is interested in an integrated approach, responses from teams with complementary areas of expertise are encouraged.

Responses have the following formatting requirements:

1. A one page cover sheet that identifies the title, organization(s), respondent's technical and administrative points of contact - including names, addresses, phone and fax numbers, and email addresses of all co-authors, and clearly indicating its association with RFI-17-01;
2. A substantive, focused, one-half page executive summary;
3. A description (limited to maximum of 10 pages in minimum 12 point Times New Roman font, appropriate for single-sided, single-spaced 8.5 by 11 inch paper, with 1-inch margins) of the technical challenges and suggested approach(es);
4. A list of citations (any significant claims or reports of success must be accompanied by citations, and reference material MUST be attached);
5. Optionally, a single overview briefing chart graphically depicting the key ideas.

Submission Instructions to Respondents

Responses to this RFI are due no later than 4:00p.m., Eastern Standard Time on 16 December 2016. All submissions must be electronically submitted to dni-iarpa-rfi-17-01@iarpa.gov as a PDF document. Inquiries to this RFI must be submitted to dni-iarpa-rfi-17-01@iarpa.gov. Do not send questions with proprietary content. No telephone inquiries will be accepted.

Disclaimers and Important Notes

This is an RFI issued solely for information and planning purposes and does not constitute a solicitation. Respondents are advised that IARPA is under no obligation to acknowledge receipt of the information received, or provide feedback to respondents with respect to any information submitted under this RFI.

Responses to this notice are not offers and cannot be accepted by the Government to form a binding contract. Respondents are solely responsible for all expenses associated with responding to this RFI. IARPA will not provide reimbursement for costs incurred in responding to this RFI. It is the respondent's responsibility to ensure that the submitted material has been approved for public release by the information owner.

The Government does not intend to award a contract on the basis of this RFI or to otherwise pay for the information solicited, nor is the Government obligated to issue a solicitation based on responses received. Neither proprietary nor classified concepts or information should be included in the submittal. Input on technical aspects of the responses may be solicited by IARPA from non-Government consultants/experts who are bound by appropriate non-disclosure requirements.

Contracting Office Address:

Office of the Director of National Intelligence
Intelligence Advanced Research Projects Activity
Washington, District of Columbia 20511
United States

Primary Point of Contact:

Chris Boehnen, Ph.D.
Intelligence Advanced Research Projects Activity
dni-iarpa-rfi-17-01@iarpa.gov