



TITLE: **Final Call for Proposals
on Digital Media Storage on DNA Support**

SOURCE: Requirements Subgroup

PROJECT: JPEG DNA Exploration

STATUS: Final

REQUESTED ACTION: For information

DISTRIBUTION: Public

Contact:
ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch

Table of Contents

● Summary	3
1. Introduction	4
2. Scope	5
3. Use Cases and Requirements	5
4. List of Biochemical Coding Constraints	5
5. Timeline	6
6. Submission Composition and Proposals Requirements	7
<i>a. Submission Elements</i>	7
<i>Submission Registration and Delivery</i>	7
<i>b. IPR Conditions (ISO/IEC Directives)</i>	7
<i>c. Contribution to Standardization</i>	8
7. Evaluation Conditions and Processes	8
8. Royalty-free Goal	8
9. Participation and Contacts	9
● References	9
● Annex A – Pre-registration form	10

Final Call for Proposals on Digital Media Storage on DNA Support

- **Summary**

JPEG DNA has been established as an exploration activity within the JPEG Committee to study use cases, identify requirements, and assess state of the art in DNA storage for the purpose of image archival using DNA in order to launch a standardization activity. To this end, multiple workshops were organized to which stakeholders were invited, use cases and requirements for the latter were identified, and several exploration studies were carried out to make a preliminary assessment of the state of the art, and a JPEG DNA Benchmarking Codec was implemented, tested and compared to a number of state of the art codecs.

JPEG standards have been successfully used in the storage and archival of digital pictures. This puts the JPEG Committee in an excellent position to address the challenges of DNA-based storage by proposing an efficient image coding format to create artificial DNA molecules.

The scope of JPEG DNA is the creation of a standard for efficient coding of images that considers biochemical constraints and offers robustness to noise introduced by the different stages of the storage process that is based on DNA synthetic polymers.

This document contains the Final Call for Proposals (CfP) on JPEG DNA issued at the 99th JPEG meeting which was held online, from 24 to 28 April 2023. Two main approaches are considered in this call:

1. Transcoding from an already existing compressed image
2. Coding an uncompressed image from scratch

Proponents are invited to express their interest in submitting technologies to be evaluated and potentially included in the JPEG DNA, to pre-register by sending the pre-registration form given in Annex A to the JPEG DNA contact (see section 9) no later than 10 July 2023 and submit a proposal package by 2 October 2023 followed by a submission of a technical description of their proposal by 23 October 2023. All proponents are expected to present their submissions and take part in discussions during an online meeting that will be held on 28-29 October 2023.

Final Call for Proposals on Digital Media Storage on DNA Support

1. Introduction

JPEG standards have been used in the storage and archival of digital pictures as well as moving images. The most popular format for storage and archival of digital pictures is the popular legacy JPEG format as described in ISO/IEC 10918 and, in particular, in parts 1, 3, and 5 of the latter standards.

While the legacy JPEG format is widely used for photo storage in SD cards, as well as archival of pictures by consumers, JPEG 2000 as described in ISO/IEC 15444 is used in many archival applications, notably for the preservation of cultural heritage in the form of visual data as pictures and video in digital format. Examples include the Library of Congress, Library and Archives Canada, Chronicling America website, and the Google Library Project. Because of its use in digital cinema, JPEG 2000 is also used for archival of movies in digital form.

In terms of technology, both legacy JPEG and JPEG 2000 formats are based on a transform-quantization-entropy coding pipeline with JPEG using the Discrete Cosine Transform (DCT) and JPEG 2000 using the Discrete Wavelet Transform (DWT), followed by quantization, coefficient reordering, and entropy coding. The legacy JPEG format has been extended to define JPEG XT, as described in ISO/IEC 18477, to include features attractive for archival applications such as lossless coding, while being backward compatible with the popular legacy JPEG format.

The latest JPEG image coding format called JPEG XL, as described in ISO/IEC 18181, also offers a number of attractive features important to archival applications, such as lossless compression and lossless transcoding from legacy JPEG to JPEG XL, resulting in smaller file sizes without numerical loss in the pixel values.

The ongoing effort in JPEG AI to produce a learning-based image coding standard is yet another potential tool that can be used in archival where the content is stored in its original form, but any post-processing such as denoising, super-resolution and enhancements are carried out without impacting the recorded content.

The purpose of this call is to invite proposals in the form of potential technologies that can be used as the starting point of an international standard for coding of digital media images that can best cope with requirements of potential applications of digital media storage on synthetic DNA support.

2. Scope

The scope of JPEG DNA is the creation of a standard for efficient coding of images that considers biochemical constraints and offers robustness to noise introduced by the different stages of the storage process that is based on DNA synthetic polymers.

Two main approaches are considered in this call:

- Transcoding from an existing already compressed image
- Coding an uncompressed image from scratch

3. Use Cases and Requirements

The following use cases have been identified as potential applications that can benefit from a coding standard for storage on DNA support. The corresponding requirements and more details about each use case can be found in the following document, which is accessible from the JPEG website: ISO/IEC JTC 1/SC29/WG1 N100252, REQ "Use Cases and Requirements for DNA-based Media Storage v1.0", 96th Meeting, Online, July 2022 [1].

- Long-term media archives and cultural heritage preservation
- Social networks cold media storage
- Preservation of medical images
- Preservation of large-scale repositories of biomedical data: beyond local data storage
- DNA coding for traceability

4. List of Biochemical Coding Constraints

The following biochemical coding constraints must be respected for any proposal submitted to this call:

- Strand length limitations
- Homopolymer runs
- GC content balance
- Repetition of patterns

More details about the definition of each constraint and their corresponding parameters are available from the following document which is accessible from the JPEG website: ISO/IEC JTC 1/SC29/WG1N100517, JPEG DNA Common Test Conditions version 2.0, 99th JPEG Meeting, Online, 24-28 April 2023 [2].

5. Timeline

The following table summarizes the main milestones regarding the timeline of this call for proposals.

98 th meeting 14-20 January 2023	Draft CfP
---	-----------

99 th meeting 22-28 April 2023	Final CfP
10 July 2023	Evaluation of anchors
10 July 2023	Pre-registration
100 th meeting 15-21 July 2023	Report on evaluation of anchors + suppl. info for proponents
2 October 2023	Submission deadline for proposals packages by proponents
6-27 October 2023	Evaluation of proposals
23 October 2023	Submission deadline for proposals technical descriptions
28-29 October 2023	Presentation of proposals by proponents
101 st meeting 30 Oct-3 Nov 2023	Report of the evaluation of proposals
102 nd meeting January 2024	Working Draft (WD)
103 rd meeting July 2024	Committee Draft (CD)
104 th meeting October 2024	Draft International Standard (DIS)
2025	International Standard (IS)

6. Submission Composition and Proposals Requirements

a. Submission Elements

Proponents are asked to submit a detailed technical description of the entire image codec, encoder implementations in software, decoder implementation in source code and the decoded test images. In addition, it is expected that support is provided to allow cross-checking of the results by independent parties when evaluating proposals. The description shall be in either PDF or Word format. Providing insufficient details might jeopardize the selection of the proposed technology for the collaborative phase of the standardization process.

Participants are required to submit materials to validate the performance of their submission according to the procedure outlined below, notably:

- Coded stream of all compressed images. The code stream must respect the biochemical constraints defined in [2]. In addition, it must include all information in order to decode the image using the decoder provided by the proponent. Furthermore, it must allow decoding any specific image from a mixture of oligos stored in a DNA support.
- Decoded images in PNG format according to the convention that will be provided to registered proponents.
- Executable encoder as well as executable and source code for decoder with instructions (compiler version, platform, libraries, etc.) on how to run for cross-checking purposes.
- Excel sheet of rate-distortion information based on the template that will be provided to registered proponents.
- Detailed description of the codec, including algorithmic description, how the biochemical constraints in Section 4 have been coped with, as well as any further details such as dependency on potential training for codecs relying on machine learning as well as complexity assessment figures such as minimum, maximum and average run times for encoding and decoding of test images by specifying the platform used. Proponents should describe whether or not their proposal is agnostic to synthesis/sequencing.

Submission Registration and Delivery

- Proponents must pre-register by **10 July 2023** by sending an email to the contacts in Section 9 of this document and filling in the form in Annex A.
- Shortly after pre-registration the proponents receive a proponent code for the purpose of anonymization along with instructions on where to access the test images for coding/transcoding, the anchor(s) software, the rate-distortion performance of the anchors in an excel sheet, as well as software to compute quality metrics.
- The materials requested in Section 6.1 should be uploaded on an FTP server which will be communicated to registered proponents no later than **2 October 2023**.

b. IPR Conditions (ISO/IEC Directives)

Proponents are advised that this call is being made in the framework and subject to the common patent policy of ITU-T/ITU-R/ISO/IEC and other established policies of these standardization organizations. The persons

named below as contacts can assist potential submitters in identifying the relevant policy information.

c. Contribution to Standardization

Proponents are informed that based on the submitted proposals, a standard specification will be created. If they submit a proposal and (part of) the proposed technology is accepted for inclusion in the standard, they will hence have to attend subsequent JPEG meetings and contribute to the creation of the different standard documents. Within this process, evolution and changes are possible as several technologies may be combined to obtain a better-performing solution.

If a proposal is selected fully or partially to be considered for the standard, its proponents agree to provide software for the part considered for inclusion in the verification model leading to the standard.

7. Evaluation Conditions and Processes

The JPEG DNA Common Test Conditions (CTC) document [2] defines the test dataset, benchmarking codecs, coding conditions (especially target rates), coding constraints, and a set of reliable objective quality metrics and subjective assessment procedures. The JPEG DNA CTC allows for an exhaustive evaluation of multiple aspects of the proposed JPEG DNA image codecs and to fully understand their strengths and weaknesses, notably regarding the proposed anchor codecs. This document was released at the 99th meeting (April 2023) as an output document.

JPEG plans to select technologies to be included in the JPEG DNA standard based on satisfying the JPEG DNA requirements [1] as well as compliance with the JPEG DNA CTC [2], such as target rates, and coding constraints.

The evaluation will be based on the results obtained through the evaluation conditions and procedures documented in the JPEG DNA CTC [2]. Several criteria and metrics will be used for the evaluation of submissions and the decision-making process:

- Validation of the quaternary coding constraints. Those proposals that do not comply with constraints may be excluded.
- Subjective quality evaluation of the proposed submission according to the JPEG DNA CTC [2]. The subjective quality evaluation results will be the primary attribute for the decision-making process for those codecs that have been validated, as highlighted in the above point.
- Objective quality evaluation with the quality metrics defined in the JPEG DNA CTC [2].
- Complexity evaluation of both encoding and decoding process.

8. Royalty-free Goal

The royalty-free patent licensing commitments made by contributors to previous standards, e.g. JPEG 2000 Part 1, have arguably been instrumental to their success. JPEG expects that similar commitments would be helpful for the adoption of a JPEG DNA standard.

9. Participation and Contacts

Proponents must register to the mailing list of the JPEG DNA AHG and should attend its online meetings.

To subscribe to the mailing list, please visit <http://listregistration.jpeg.org> or in case of problems contact lists@jpeg.org

Contact list:

Fernando Pereira (JPEG Requirements Subgroup Chair)

Email: fp@lx.it.pt

Thomas Richter (JPEG Image Coding and Quality Subgroup Chair)

Email: thomas.richter@iis.fraunhofer.de

● **References**

- [1] ISO/IEC JTC 1/SC29/N100252, Use Cases and Requirements for DNA-based Media Storage version 1.0, 96th Meeting, Online, 25-29 July 2022
- [2] ISO/IEC JTC 1/SC29/WG1N100517, JPEG DNA Common Test Conditions version 2.0, 99th JPEG Meeting, Online, 24-28 April 2023

- **Annex A – Pre-registration form**

Proponents must pre-register by **10 July 2023** by sending an email to the contacts in Section 9 of this document and filling in the form below.

Title of the Submission:

.....

Acronym of the Submission

.....

Category of the Submission

Check one or both depending on the category you are submitting to.

Transcoding	
Coding	

Contact(s)

If several organizations are involved, provide at least one contact per organization.

First name, Last name:

.....

Email Address:

.....

Affiliation and postal address:

.....

(repeat new additional contacts as many times as needed)