# ELFT-EFS Evaluation #1

# Questions and answers to date

*Last updated 21 August 2009*

**Q1: Regarding field 9.308 (RQM): 6 valid values (quality codes) are defined (0 through 5) but only 2 bits are allocated, which are not enough for representation of 6 values. How would this be resolved?**
> A1: It appears that you are referring to the EFS draft 0.2 or earlier, which had the issue you mention for one format. In EFS 0.3 and 0.4, field 9.308 can use the formats specified infield 9.309, Uncompressed or Run-Length Encoded. The data for the ELFT-EFS evaluation will use the Uncompressed format for simplicity, which uses one character per grid cell. The current specification can be obtained from http://fingerprint.nist.gov/standard/cdeffs/

**Q2: We are not sure about instructions for multiple-encounters. The instructions given are to keep track of multiple-encounters during enrollment, but it is unclear what to do during matching: Are we free to use this info in any manner we choose? Or should we ignore this information for now and treat multiple-encounters as if they were different subjects? Or, will there be a flag specified in the API telling us which course of action to follow?**
> A2: In Evaluation #1, whenever different fingerprint sets belong to the same person, that will be communicated to the participants at the time of enrollment, and the vendors can and should take full advantage of that.
> On the candidate lists for Evaluation #1, the subject and finger position are designated. The candidate list does not indicate whether that was due to a match against a specific roll or slap, or a combination.
> For a given test, all gallery subjects (mates and non-mates) should be expected to have the same types and numbers of exemplar sets. For example, the general case (the gallery on which most tests will be run) will have 10 rolls + 10 (presegmented) slaps per subject.

**Q3: Regarding Sequential and Multithreaded, I understand that search workload is critical in this evaluation. Therefore, I plan to prioritize multithreaded method for Search function than Exemplar Enrollment. As you know, time for Latent Enrollment is negligible. Is it acceptable for you to have Multithreaded function for search (matching) and Sequential function for enrollment?**
> A3: Yes.

**Q4: Please let us know if latent data for Evaluation #1 have skeleton or ridge flow data. Although all data for Public Challenge have skeleton and ridge flow data which were created by skeleton, it is doubtful to have such data for Evaluation #1. Because skeleton and ridge flow data have very rich features compared to other features and our strategy will be focused on using such data if exist.**
> A4: Some but not all of the latents will have skeleton/flow data.

**Q5: You 'divide and conquer' over the probes, but not over the gallery. Is there a way to break-up the gallery over multiple processes in the current design? This will allow the use of larger templates. Ideally you would allow a vendor to run 2 processes - each working on a portion of the gallery, and then you would collate the top 100 results from those processes to generate a top 100. If you do not provide that ability, are the SDK providers allowed to provide such a solution?**
> A5: Perhaps for future tests, but right now this presents numerous other problems due to complexity.

**Q6: In the public challenge multiple exemplars per subject were expected to be returned as individual candidates by the matcher. Will this be the case for the Evaluation?**
> A6: (see also Q2) For Evaluation#1, full use may be made of all exemplars available per subject. The candidate list will indicate the subject ID and the finger position, even when

rolled and slap (or multiple exemplar sets) are available for the subject. Thus, if so desired the candidate returned and its associated score may represent a fusion of all available exemplars for that subject.

**Q7: If the probe/latent feature set is different for different searches, what do we assume in determining the throughput? (Different parts of the feature set have different computation time.)**

A7: For throughput timing, assume that the feature set is identical to that present for the public challenge subtest L3: Baseline EFS features, including image. The spreadsheet ELFT-EFS_Fields.xls lists the fields available. (See also Q53)

**Q8: Do we have to implement all of the EFS (Extended Feature Set) 0.4 specification for Evaluation #1? Will we have to implement both 9.309 formats for the Quality Map?**

A8: All of the data in Evaluation #1 will be limited to the fields and formats included in the public challenge subtest L3. Therefore, field 9.309 will be set to "UNC" for all of the Evaluation #1 data as well.

**Q9: In the Public Challenge dataset, a few of the cells for the ridge flow (9.310) seem to be off. An example of marked extended data with 'off' value: L3\L002G_JURIED_NP.lffs, Row 54 , 25th block orientation is 0x5A=90, and 26th block orientation is 0x16=22. The difference of these two angles is 90-22=68. It is in an area where ridges go smoothly. There is no data for the 24th block while the orientation for the 23th block is 45.**

A9: The ridge flow values were extracted from the ridge tracing. The tracing in the cell mentioned in your example has a small downward bend, which resulted in the ridge flow angle you note. The algorithm currently used is based on the current cell direction, without using a larger window, which results in sensitivity to local changes. In discussions among CDEFFS (in 2006), the committee considered and rejected recommending/requiring a specific window size for determining ridge flow. One implication of using a small window size is that the consumer of the ridge flow data may choose to blur the data (effectively applying a larger window after the fact), but has the option of choosing the window size to use.

**Q10: Our understanding is that a technology provider can participate in public challenge but opt to not participate in evaluation #1. Is this correct? If so, does the provider have to inform NIST about their withdrawing before a certain deadline to remain anonymous?**

A10: You may participate in the Public Challenge and not participate in Evaluation #1. Public challenge anonymity is not related to participation in Evaluation #1.

**Q11: Can the storage device be used during matching? Being able to use the hard #drive(s) would compensate for the lack of addressable memory. Knowing how many #cores share the same hard drive will also help optimizing available resources.**

A11: Yes, the enrolled gallery will be readable from local storage. As many as all 4 cores shall have access to the enrolled gallery. We assume that a single multithreaded SDK process with 4 threads (one per core) will more efficiently be able to access storage than 4 separate sequential SDK processes.

Multithreaded and sequential SDKs will be executed with access to a single copy of the enrolled gallery on the local storage device (shared by all cores).

**Q12: How much local storage will be available to the matcher processes?**

A12: We are assuming that for the largest gallery (100k rolls+slaps), 150Gb per blade will be sufficient. Please respond if this is not reasonable.

**Q13: I understand probes will be searched by batch (=subset). What is the typical #size of a subset?**

A13: The baseline set of latent probes contains approximately 1500 latents. If the submitted SDK is multithreaded, we will divide the number of latents in the sub-test among N instances of the SDK executing on N blade servers. For example, if 10 blade servers are available to test a single SDK, 10 instances of the SDK will be executed, each with a subset of 150 latents.

**Q14: Are you expecting SDKs to implement multithread by searching several probes in parallel?**

A14: We have no preference on how the multithreaded SDK makes use of the 4 cores per blade server. As you say it could search 4 latents in parallel. However it could also parallelize each (single) latent search across all 4 cores. The implementation details are left to the SDK provider.

**Q15: Will finger position be indicated for the latents?**

A15: No. Finger position is set to 00 (Unknown) for all latents.

**Q16: With respect to the comment "Not all features will be included in all searches." Can you please clarify if this includes the possibility that the EFS markups could be incomplete?**

A16: The spreadsheet ELFT-EFS_Fields.xls lists the PROPOSED field combinations for Evaluation #1. Please respond if these are not reasonable.

**Q17: Our understanding is that the data in fields 9.310 and 9.372 are not marked manually but generated by software. Is this correct? If yes, (1) can you please let us know what software is used to generate the data? and (2) Are there any other fields that are not marked manually?**

A17: The skeletonized image (9.372) WAS marked manually. The ridge flow map (9.310) was generated from the skeleton using the ULW libraries — the ridge flow map is the only field that was generated by software rather than direct human markup.

**Q18: It appears that the throughput timings based only on the image+all EFS type of test. I don't know if that is your intention, but I would expect that the image-only search for most vendors is the slowest.**

A18: Throughput timing will be based on Subtest LE (Baseline EFS features, including image). We recognize that for some implementations, throughput for image-only searches (Subtest LA) may be slower due to less effective screening. We assume that throughput for image only searches (Subtest LA) will be slower by a factor of up to 2x.

**Q19: It is not clear from the table if the latent images will be available for all of subsets or if some of the tests will be probes will be run without the image as in the public challenge.**

A19: Subtest LD (IAFIS/ EFTS equivalent) will be run both with and without the image. All other subtests include the image. The spreadsheet has been revised to make this clear, adding Subtest LG (IAFIS/ EFTS equivalent without image).

**Q20: Regarding the WSQ implementation for the exemplars. Some encoders do not follow the norm strictly and this could pose decompression issues during the test. I do not believe WSQ compliance should have an impact on accuracy in this benchmark. Would it be possible to send an simple executable or library that checks if all exemplar files are decompressed correctly (ie a OK / KO output for each record)? Such a test could be run early on this month in order to solve any WSQ issue before August.**

NIST is providing a library that SDKs shall use in order to extract all fingerprint images (WSQ compressed) from 10-print exemplar (AN2) files. The library is built from our NBIS library which contains our own WSQ decompression code. Prior to testing we will ensure that all WSQ exemplar print images used in the test are retrievable using our library. The library will contain a function that takes as input an AN2 file pathname, and outputs a list of pointers to all images contained within the AN2 file in uncompressed (RAW) format (it also will return the height, width, impression type, resolution and finger position of each image). The July 10 version of the ELFT-EFS Evaluation Test Plan defines the interface to this library.

**Q21: Is it permissible to submit our extractor SDKs for Linux, but our latent search SDK for Windows?**

Yes, each individual SDK may run on any of the test platforms specified in the ELFT-EFS Evaluation Test Plan.

**Q22: According to the test plan, "E1 . 100,00 subjects; 10 rolled & 10 plain impressions." I understand that E1 consists of 200,000 ANSI/NIST files. Am I correct?**

>Yes.

**Q23: Regarding the timing requirements, the test plan states: "0.05 sec/exemplar set (20 exemplar sets/sec, per latent, assuming an exemplar set consists of 10 rolled and 10 segmented slap fingerprints)." Does this mean 0.025 sec/exemplar set consists of 10 rolled (E2) OR 10 segmented slap fingerprints (E3)"? Or it is OK to spend up to "0.05 sec/exemplar set for E2 or E3?**

>The maximum average latent search time for sets LC-LG is 0.025 sec/10-finger exemplar set, per latent. These timing requirements have been restated more clearly in terms of 10-print exemplar sets in the July 10 version of the ELFT-EFS Evaluation Test Plan. Therefore, the permissible latent search time per subject will scale linearly with the number of 10-finger exemplar sets per subject (e.g. 0.05 sec/subject for E1, since each subject has 2 associated 10-finger exemplar sets, and each set requires 0.025 sec)

**Q24: Regarding the timing requirements, the test plan states: "E6 . 10,000 subjects; 4 sets of 10 rolled impressions." Does this mean 0.025 sec per "4 sets of 10 rolled impressions"? Or is it OK to spend up to 0.1 sec per "4 sets of 10 rolled impressions"?**

>Yes, 0.1 sec/subject is permissible for E6 because each subject has 4 associated 10-finger exemplar sets. See our answer to Q57 for a more complete answer.

**Q25: Is it OK to have make single template per subject? For example, we would like to create a single template for E1 exemplars (from rolled ANSI/NIST file and slap ANSI/NIST file). Or are we required to create separate template per ANSI/NIST file as shown below:**

**"E.g. if exemplarFeatFilenames = ["/input/path/E9999_1.feat" , "/input/path/E12345_1.feat" , "/input/path/E12345_2.feat"] and galleryDirectory= "/this/gallery/", the software inserts the exemplar feature sets into the stated gallery, associating the two files with the subject number 12345."**

>The enrollment process (i.e. create_gallery()) may merge templates created by the extraction process (i.e. extract_exemplar()). We don't regulate the format of the enrolled gallery. We leave this up to the SDK provider, so if you want to merge all templates (ie. ".feat" files) for a subject into a single template in the gallery that's OK.

**Q26: The test plan states "Recognition rank ranges from 1 to 50, as 50 is the (maximum) candidate list size specified in the API " but then later says that the candidate will be top 100, as was the Public Challenge. Can we get clarification?**

>The required candidate list length is 100. The July 2 version of the ELFT-EFS Evaluation Test Plan had a error (i.e. "1 to 50, as 50" should've been "1 to 100, as 100") which has now been corrected in the July 10 version.

**Q27: I have a question regarding the functionality of the API create_gallery(). What exactly does it mean by "associating exemplar feature sets that share the same ID"? Could you please elaborate on this? I'm assuiming it has something to do with when outputing candidate lists in which only the IDs and finger numbers are to be listed, but I just can't see how this comes into play.**

It simply means that they should be stored in the gallery in some way so that the matcher will know which exemplar sets belong to a specific subject. This is so that latent_search() can use all available exemplar sets for each subject when producing a candidate list.

For example to create gallery E1, the function create_gallery() will be input with a list of 200,000 extracted 10-finger exemplar feature set filenames (each feature set was produced by previously calling extract_exemplar()). The resulting gallery will then contain 100,000 subjects records each with two associated 10-finger exemplar feature sets (which may be stored separately, or combined in some way).

In the example above, the candidate list will specify the subject and finger number, but not which of the subject's two 10-finger exemplar set(s) were used to produce the score. The score for a given candidate may be produced by using any or all (e.g. by fusion) of a subject's associated 10-finger exemplar sets.

**Q28: Regarding gallery dataset sizes, it is stated in section 3.8 that the largest size gallery used for Evaluation #1 will contain 100,000 subjects having two 10-finger exemplar sets (rolled and plain impressions) per subject.**

**(1) All exemplar subsets defined in section 4.2 are datasets of 10,000 subjects. Is E1 supposed to be the largest gallery dataset with 100,000 subjects?**

**(2) Does '100,000 subjects' mean 100,000 UNIQUE individuals?**

**(3) If the answer to question (2) is no, what is the number of unique individuals for each gallery datasets (E1 through E6)?**

**(4) Can you please provide the follow numbers for E1?**

**a) total number of exemplar sets with both rolled and plain impression,**

**b) total number of exemplar sets with rolled impression alone**

**c) total number of exemplar sets with plain impression alone**

**(5) Our understanding is that E4 contains 2*10000 = 20000 sets of 10-finger exemplar data, and 30,000 10-finger exemplar sets in E5, 40000 10-finger exemplar sets in E6. Is this correct?**

    (1) E1 will have 100,000 subjects; whereas E2-E6 have 10,000 subjects. The July 2 version of the ELFT-EFS Evaluation Test Plan had an error in section 4.2 for E1 (i.e. it stated "E1 - 100,00 subjects" instead of "E1 - 100,000 subjects). This has been corrected in the July 10 version.

    (2) Yes.

    (3) skipped

    (4) All exemplar sets will be 10-finger exemplar sets. For E1, there will be 100,000 10-finger rolled exemplar sets, and 100,000 10-finger plain exemplar sets, for a grand total of 200,000 10-finger exemplar sets.

    (5) Yes.

**Q29: Regarding multiple sets of exemplars for one individual,**

**(1) What is the maximum number of exemplar sets associated with one person?**

**(2) "Exemplars that are known to belong to the same person will always share the same subject ID". What is the naming convention for subject ID? Is it fixed length? What is the length/maximum length of the subject ID?**

**(3) Section 3.6 indicates that the association between exemplars for same subject "will be made explicit in the exemplar enrollment stage: at the time of enrollment, " Does this cover all exemplar sets?**

    (1) Four 10-finger exemplar sets, as specified for exemplar subtest E6.

(2) The 'subjectID' is 6 digits (fix-length, leading zeros) as specified by the July 10 version of the ELFT-EFS Evaluation Test Plan (see section 5.4.4b).

(3) Yes.


**Q30: Regarding the API, the notes for function 'create_gallery()' indicates "The exemplar feature filename will be formatted "E"subjectID "_" instance ".feat", where subjectID is the numeric ID for the subject ".   The notes for function 'extract_exemplar()'indicates that the function should take exemplar data file exemplarFilename.an2 and output feature data file exemplarFilename.feat. It also indicates that the .an2 file will contain either 10 rolled or 10 segmented slap images.**

**(1)  The input .an2 file won't contain both rolled and plain fingerprints, is this correct?**

**(2)  The name of the .an2 file is formatted as "E"subjectID "_" instance ".an2. Is this correct?**

**(3)  For subject with both rolled and plain datasets, how is the 'instance' formatted to indicate that the two files are the rolled and plain datasets for same subject?**

**(4)  What is the maximum length of file name?**

**(5)  What is the maximum length of directory name (including path)?**

(1) Correct.  Each .an2 file for the exemplars will contain only 10 fingerprints (either all 10 rolled or all 10 plain)

(2) See section 5.4.4b (2$^{nd}$ to last paragraph) of the July 10 version of the ELFT-EFS Evaluation Test Plan for the format of the exemplar filename.

(3) The 'subjectID' will be the same for the 10-finger rolled exemplar set and the 10-finger plain exemplar set which belong to the same subject.  The value of 'instance' in this case will differ between the two sets.  The precise value of 'instance' for either set is TBD and not important since metadata in the related ANSI/NIST files (which is returned by extract_image_data(), see sec. 5.4.3a) will indicate the impression type of the prints (i.e. rolled or plain).

(4) All pathnames (which includes the filename portion when it exists) used by the API have a maximum length of  255 characters.

(5) See above answer.


**Q31: Regarding the API extract_latent() function, the 'example' says "create /out/path/L12ABC.an2". Do you mean create /out/path/L12ABC.feat ?**
Yes, this was an error in the July 2 version of the ELFT-EFS Evaluation Test Plan which as been corrected in the July 10 version.


**Q32: Will there be any legal document to protect the participant's IP?**
Please refer to sections 6.4, 6.9 and 6.10 of "Application to Participate in the Evaluation 1 of Latent Fingerprint Technologies Extended Feature Sets (ELFT-EFS Evaluation 1)" for language relevant to software ownership and usage restrictions.

Only NIST personnel are involved in the execution of software under test (which is executed on dedicated isolated NIST test hardware).  We use a variety of logical and physical means to protect all software in our test environment (starting from the moment you encrypt it and send it to us), affording it the same protections used to safeguard our test data which is Sensitive But Unclassified (SBU) U.S.  government use only.  We have been testing "SDKs" since 2003:
http://fingerprint.nist.gov/PFT/index.html
http://fingerprint.nist.gov/slapseg04/index.html
http://fingerprint.nist.gov/slapsegII/index.html
http://fingerprint.nist.gov/minex04/index.html
http://fingerprint.nist.gov/minex/Status.html

**Q33:  For our submission, apart from having proprietary DLLS and the "libelftsSDK.dll", we have configuration files. The configuration folder must be in the folder where the SDK dll is located. Please let us know that such an arrangement is acceptable.**
> Yes, that is acceptable.


**Q34: We are planning to submit a 64 bit SDK targeted on using the 64 bit OS specified in the Test Plan (at least for matching).  In the past, NIST has required software to be cygwin "compliant". But there does not exist any 64 bit cygwin test platform. Will you be using 64 bit Windows test harness for such SDKs or there is some alternative arrangement that we should be aware of?**
> We're currently evaluating "MinGW-w64" (which may be run under Cygwin, which is native 32-bit) as a means to build a 64-bit test harness.  If this is not a workable solution then our fallback will be the Microsoft C compiler.  So long at the submitted libraries have "C" (not C++) entry points, either solution should work.


**Q35: Can you provide any guidance on acceptable time for implementing the create_gallery() function? We are seeing that it takes us a few ms per finger.**
> 100ms per ten-print is the upper bound.


**Q36: The test plan document mentioned that they will be using a C test app to link to our SDK. We assume we will need extern "C" around the extract_exemplar() (and other SDK functions). Can you confirm?**
> Yes, the test harness will be C code, and thus a C interface is required.


**Q37: I understand that most of latent prints have effective pattern type data (not unknown). Please let us know estimated proportion of unknown pattern types in latent prints of Evaluation #1 test set.**

 *>It is recognized that for some implementations, throughput for image..only searches (Subtest LA)*
*>may be slower due to less effective screening. It is allowable for throughput on Subtest LA*
*>(image only) and LB (image+ROI) to be slower by a factor of up to 2x than the stated search time.*

 **Note: This is true when most latent prints have effective pattern type data.**
> The proportion of latents with the pattern type undefined will be approximately the same as was true for the public challenge data.


**Q38: Can we assume all "Manually Entered Data" such as Pattern types, Orientation , etc. are reasonably reliable? Or do we need to consider "manual mistakes"?  For example, how do you evaluate SDK when you find incorrect pattern type?**
> The latents for evaluation #1 are being marked up in TWO ways:
> A) The latents are marked to the best of an individual examiner's ability, using the latent alone (as was done in the public challenge dataset)
> B) The latents are also "ideal groundtruthed" using the exemplar in addition to the latent (as was done in the original SD27)
> A shows the results achievable using real world markup, but includes the possibility of examiner error
> B is not possible in the real world, but shows ideal performance and allows comparison of matchers without the confounding effects of examiner error

**Q39: Matching time against Rolled and Plain prints**

*>0.025 sec/10-finger exemplar set*
*>Rate of 40 exemplar sets/sec, per latent (exemplar set = 10 all rolled or all plain prints)*

Generally speaking, the matching against "10 all plain prints" may be faster than one against "10 all rolled prints" because of smaller effective impression size.

Do we need to meet this requirement on "10 all rolled prints" (E2)? Or is it OK to meet this requirement as an average (or fusion) of "10 rolled prints" and "10 plain prints" (E1)?

> An average is acceptable.

**NEW Q40: Regarding Q38, (1) Are all the features, including EFS, 'ideal groundtruthed' (wherever it can be applied), or only pattern type and minutia? (2) Are you going to run two tests over LC through LG, one with 'real world' markup and the other with 'ideal groundtruthed' markup?**

> (1) Yes, and (2) Yes.

**NEW Q41: The test plan says: "Exemplar images will be approximately upright (in the same orientation as they were captured)." Since slap images are often rotated (especially for inked prints or small platen livescan), this seems to be a contradiction.**

> The images will be in the same orientation as they were captured. The text should read "Exemplar images will be in the same orientation as they were captured."

**NEW Q42: the Test Plan states that all exemplars will be wsq compressed. It also says that the slaps are segmented out. If they are decompressed, cropped, and recompressed, doesn't that introduce artifacts?**

> The segmented images are not recompressed. The will be stored uncompressed after being segmented.