Federal Bureau of Investigation (FBI)
Next Generation Identification (NGI)
Latent Fingerprint Search Strategies

Under the FBI's Integrated Automated Fingerprint Identification System (IAFIS), some agencies treated all latent searches the same: always marking minutiae and counting ridges, and always comparing all twenty candidates returned. With the Next Generation Identification (NGI) System, it may be more effective to think in terms of **search strategies**: deciding how to conduct latent searches based on the requirements and implications of that specific case and that specific latent. Search strategies seek to optimize tradeoffs between effectiveness (maximizing the likelihood that a search will result in a hit) and efficiency (minimizing the effort required for searching and comparing candidates).

For most agencies, different cases may have widely different requirements, based on the case priority, how much examiner time is available, forensic relevance of the latent, quality of the latent, and the workload:

- minor cases that would otherwise never be searched may justify only a minimal effort ("low-hanging fruit");
- a homicide may necessitate an exhaustive search and many times the amount of effort of a routine case ("no stone left unturned");
- backlog, cold cases or an overwhelming workload may benefit from prioritization ("biggest bang for the buck");
- routine cases need to balance among these approaches.

There are several possible strategies that could be used to accomplish different objectives when searching against NGI. The purpose of this document is to provide suggestions to latent print examiners and laboratory managers on how to make effective and efficient use of NGI latent services based on the needs of their agency. This guidance was developed through analysis of data generated from the NGI algorithm.

(*Note: this document is limited to latent fingerprints.)*

How do I select which latents to search, and how do I search them?

How do I decide whether a latent is of adequate quality to search NGI?

Lower quality latents are less likely to hit, but any latent that is of value for individualization can potentially be matched by NGI; even some latents that examiners would consider "No Value" can be matched by the NGI system. Therefore, the practices of assessing latents that are of value but low quality as "Not Suitable for IAFIS" (or establishing a distinct quality threshold for "AFIS searchable") are not supported. If you choose a part of a search strategy to focus on those latents that are more likely to hit, Latent Quality Metrics (LQMetrics) Software will estimate the probability that each latent will hit assuming the mate is in the gallery, is of sufficient quality, and has sufficient overlap with the latent. (*See accompanying LQMetrics User Guide on FBI Biospecs at https://www.fbibiospecs.cjis.gov/Latent/PrintServices for more information)*

How do I decide which latents to search first?

When faced with multiple latent fingerprints on a case or a backlog of multiple latent images, it may be cost effective to prioritize workload within or across cases by selecting which latents to search first. After taking into account case priority and probative value of a latent, LQMetrics may be used to prioritize searches, so that the latents which are most likely to result in a hit are searched first.

How do I decide whether to do an image-only Latent Fingerprint Image Search (LFIS) or to do a feature markup Latent Friction Ridge Features Search (LFFS)?

Image-only searches were found to succeed almost as often as feature markup searches. Because image-only searches require less effort to submit, this is generally the preferred approach for the first search attempt. In most cases, if a latent hits at all, it will succeed both as a LFIS and LFFS.

If resource constraints dictate that only one search will be performed, the total cost of submitting the search and comparing the candidates should be considered: while preparing an image-only search requires less effort, a feature markup search is slightly more likely to succeed. If many candidates will be compared, the additional markup search is slightly more likely to succeed. If many candidates will be compared, the additional markup effort may represent a small marginal increase in total effort.

Contrary to the expectation that poor quality latents would benefit from markup more than high quality latents, no such relation was found. Quick Minutiae searches (LFFS) resulted in slightly higher hit rates than image-only searches (LFIS) regardless of whether the latents had high or low LQMetrics scores.

How do I decide whether to search a latent multiple times?

If the first search does not hit, there is a chance that additional searches of the same latent will succeed: after both LFIS and LFFS searches have been conducted, additional LFFS searches (with distinct feature markups) can be conducted if the case priority warrants the additional effort.

For a given latent, the cumulative probability of a hit will increase incrementally as multiple searches are performed, but the marginal benefit of each subsequent search is reduced. In addition, one or more failed attempts are likely to indicate a problem that subsequent searches will not overcome (such as the mate not being in the database).

How do I review search responses effectively and efficiently?

How do I minimize effort in reviewing responses?

If the strategy is to collect low-hanging fruit (minimizing examiner time spent

reviewing responses in order to process as many latents as possible, at the cost of missing some identifications), then the amount of effort per identification can be drastically reduced by instituting rank or score thresholds. The vast majority of hits are at rank 1, and the majority of those have high matcher scores. The number of candidates compared can be limited based on a score threshold or rank threshold. For low-priority cases, this may mean only reviewing one candidate per search or (for the lowest priority cases) only reviewing responses if rank-1 candidate is above a score threshold. This will mean that possible identifications will be missed, but this may be the best solution for minor cases that might not otherwise have been searched.

How do I prioritize the responses I review?

When many searches have been submitted and resources available for performing comparisons are limited, it may be desirable to prioritize the workload by selecting responses so that the most likely identifications occur first. An examiner can select which response (candidate list) to review next by selecting the response with the highest scoring rank-1 candidate, and then either complete all comparisons for that candidate list, or by selecting the response with the next highest score (bouncing between candidate lists). Alternatively, an examiner may choose to review all rank-1 hits (across multiple responses) before reviewing the subsequent candidates.

How do I decide how many candidates to review?

The general approach is to order the candidates such that the most likely matches are compared first, then to stop comparing when the probability of success on the next comparison has diminished to the extent that the cost cannot be justified.

The stop criterion may be driven by a fixed resource limit or by a targeted cost-benefit objective. If resources are fixed, the examiner can simply compare candidates until the resources are spent (e.g., a fixed number of staff hours may

be allocated to a batch of work). Targeting cost-benefit objectives can be based on rank thresholds or score thresholds: a policy decision could be made that comparing only a specific number of top candidates (five or ten), or only candidates above some fixed score threshold is cost effective, whereas comparing other candidates is not cost effective. Different thresholds could be set for different case priorities or other circumstances.

When deciding on a stop criterion, consider the tradeoffs between efficiency and effectiveness. When a mate is returned, it is usually found at rank 1 on the candidate list. It may be practical to start with a conservative choice score threshold, then gradually adjust up or down based on an agency's experience over time.

How do I select and implement latent search strategies?

How do I decide on a latent search strategy?

The selection of an appropriate search strategy allows the examiner to make effective use of time appropriate with the search accuracy needs of a specific case: different search strategies allow tradeoffs between examiner time and the probability of making identifications, as well as prioritizing workload. Latent search strategies require policy decisions: How much additional effort is acceptable for a small gain in identification? Is a reduced identification rate justified by a substantial reduction in examiner time?

An agency may select different search strategies based on a variety of factors, including the priority of a case, forensic relevance (probative value) of the latent, quality of the latent, number of latents in the case, overall workload, and staffing availability. Each of these may affect the tradeoffs among the benefit of making an identification, the risk of missing an identification, and the appropriate amount of examiner time.

How do I minimize the effort needed for searching latents and

reviewing responses?

For low-priority cases (possibly including property crimes, backlog, or cold cases) where it is not practical to do an exhaustive search of every latent, examiner time can be reduced significantly for only a moderate reduction in identifications. This would result in a substantial increase in the number of latents that could be searched and identified.

The approaches to minimize the effort needed for searching latents include:

- Investing less effort searching each latent – Examiners can reduce search effort by submitting LFIS (image-only) searches rather than LFFS searches, and by limiting the number of times each latent is searched. LFIS searches are almost as effective as LFFS searches, but require a fraction of the examiner's time. Re-searching latents with different markup can be effective, but the chances of success drop rapidly with each search.
- Investing less effort comparing candidates – Examiners can limit the number of candidates compared based on a rank or score threshold.
- Searching fewer latents – Examiners can select fewer latents to search using LQMetrics or based on factors such as case priority, probative value, or the number of latents in the case.

For example, if there are thousands of cold cases that are impractical to search exhaustively at the present time, the latents could be searched as LFIS transactions by junior (non-examiner) staff, and examiners would only have to be involved to review the fraction of NGI responses with candidates that are above a certain score threshold indicating that they are likely to be hits.

How do I maximize the probability of an ID in an important case?

For the highest priority cases, the following approaches offer an increase in the probability of an ID, but at a substantial increase in examiner time. One should expect diminishing returns for each increment in effort after the initial search.

Approaches to maximizing the probability of an identification include"

- Multiple searches – If the initial search is not successful, multiple different LFFS searches, by the same or different examiners, can be conducted for a small gain in identifications.
- Comparing more candidates – Increasing the candidate list length and reviewing all available candidates will increase the probability of an identification – although the marginal gain, especially after the top 5 or 10 candidates, is small.
- Search latent prints of lower value – While the lowest-quality latents have a lower chance of hitting, any latent sufficient for manual comparison might hit in NGI.

How do I prioritize searches so that the most likely identifications occur first?

For some cases that involve multiple latents, it is desirable to conduct searches and review responses so that the most likely identifications occur first. In general, this involves first using the approaches to minimize examiner effort then (if necessary) following up with additional searches or more detailed reviews of the candidates.

The approaches to prioritizing workload include:

- Searching LFIS first, then LFFS – Since examiner time spent preparing LFIS searches is significantly less than for LFFS searches, it is more efficient to conduct LFIS searches first, then only conduct LFFS searches if the LFIS searches do not hit and the case warrants the additional work.
- Sorting latents by quality – Because LQMetrics estimates the likelihood of a successful NGI search, sorting searches by descending LQMetrics values frontloads those most likely to hit.
- Review rank-1 responses first, and/or sort responses by score – Reviewing the rank-1 hits across multiple responses, or sorting multiple responses by descending matcher score frontloads those responses most likely to contain hits.