

# Making a Case for Effective Latent Print Casework Metrics

George Kiebusinski, Austin Hicklin (Noblis), and Melissa Taylor (NIST)

How do laboratory managers know how effective and efficient their latent print operations are? Latent print examiners are faced with a variety of tasks beyond the ACE/V comparisons that form the core of their work, including processing and imaging physical evidence; preparing and conducting AFIS searches; reporting examination results; performing verifications, administrative reviews, and technical reviews; participating in inter-examiner consultations; preparing courtroom presentations; providing depositions or testimony; and completing administrative paperwork. Some laboratories have made concerted efforts to collect and use metrics to measure and analyze the work conducted by their latent print units. However, there is no standard or best practice for data collection, recording, or reporting of casework metrics across the latent community. The lack of data is exacerbated by variances in process flows, data characteristics, and terminology. NIST and Noblis, in cooperation with OSAC, are working with the latent print community to define a standard set of latent casework metrics, so that there is a uniform way to assess the effectiveness and efficiency of latent print operations.

## Benefits of a Defined Casework Metrics Capability

There are wide differences in how latent print unit performance is measured and reported. While some laboratories have taken steps which go a long way in addressing this issue internally, many more laboratories have no program in place for assessing their performance or effectiveness, internally or against other agencies. Latent casework metrics enable a laboratory manager or supervisor to do the following:

- Assess effectiveness and efficiency of the laboratory overall, as well as of individual examiners
- Determine the time spent on the various parts of the workflow process so that improvements can be developed
- Measure the impact of newly implemented changes to operating procedures on staffing, effectiveness, and productivity
- Optimize process flows to improve responsiveness to investigators
- Measure staff productivity to determine if corrective actions or retraining are required
- Provide data for justifying staff and/or other resource requests

Laboratory managers and supervisors who use metrics to evaluate their latent print operations have provided a number of examples how metrics have been used successfully for operational improvements:

- **Process optimization:** *Metrics helped us identify where a bottleneck occurred during our examination process. We identified that the administrative portion of our cases was the source of our backlog. In order to streamline the process we choose to adopt a standard form for case notes. This minimizes any variability between technical examiners and how much time an admin reviewer takes to go over the report.*
- **Workload management:** *From the information tracked in our LIMS we synthesize monthly statistics, and evaluate productivity of our staff. We can determine what our monthly output will be based on how much is in our backlog, and our lab's productivity at any given time. We also use this information to argue for new positions. Lastly, we determine what extra activities our examiners are participating in based on their output.*
- **Examiner productivity:** *A few years ago we decided to take a look at each examiner, assessing how many cases they handled, how many they decided were AFIS latent quality, and how many of those latents hit in AFIS. We determined that some examiners have really high hit rates, but that was inversely proportional to how many AFIS searches they conducted: they only put the best*



*stuff in and are getting a high hit rate, as compared to other examiners that search a lot of latents and have a very low hit rate. We were able to establish who our AFIS superstars are: the examiners who submit a fair amount of stuff, but got good hit rates. These are the people that do it really well, and we wanted our other examiners to start modeling their AFIS input like the others.*

- **Examiner productivity:** *We started collecting metrics for examiner performance. There seems to be an effect, that just by knowing that we are collecting data, examiners who have been performing poorly have been performing better. So we just started collecting a bunch of different stuff, and just comparing everyone against everyone else. We're still fine tuning it – we have the data, but the question is what the best way to interpret it is? Our backlog is going down fairly quickly now.*
- **Workload management:** *We found that collecting information based only on the number of cases did not work. We collected information on how many cases we worked and how many latents were in each case, and found that over the years the number of latents per*

case has increased, in part because the proportion of latents you can search through AFIS has skyrocketed. In particular, turnaround time needs to consider the number of latents per case, not just for cases overall.

- **Examiner productivity:** We track monthly stats of our people to determine performance, and understand which staff are meeting expectations and balancing other career goals. We use this to help determine who is permitted to work overtime, and who is allowed to do research projects.
- **Workload management and Examiner Productivity:** Good metrics helped demonstrate that caseload exceeded capacity, even though we could show that we were more efficient than other labs. We could show that we could not improve turnaround time and throughput without more staff. Inefficiency from turnover resulted in 10-20% loss in productivity during the training period. Data showed that low salaries were correlated with turnover. Resulted

backlog: in particular we had a huge backlog of lower priority cases. We looked at our metrics: how much work we are doing, how many latents we were getting every month, how many comparisons, and the AFIS candidate list size. We were looking at 20 candidates on the search list and we did a cost benefit analysis of what the utility was for looking at all 20 rather than looking at 10. We discovered that getting a hit at ranks 11-20 took far more comparisons than getting a hit at ranks 1-10. This is because looking at ranks 11-20 doubles the effort for every search that does not result in a hit. So we made a policy change to only look at the top 10 candidates. Our hit rate (per search) dropped only a little, but the total number of hits increased substantially because we were able to complete many more searches.

These success stories underscore the benefits of collecting and using metrics to improve latent print operations. However, we have found that these results are difficult to compare among laboratories, because there is so much variability in practices, terminology, and in what data is being collected.

It would be advantageous to the latent print community to develop a metrics program that is practical to implement, easy to understand, and useful to the management of latent agencies. By providing a standard set of metrics, latent processes can be optimized to provide higher levels of performance and provide a basis for improved service levels that can be compared across agencies.

Defining terms and identifying metrics that are acceptable to the latent print community is a fundamental step in the development of more effective business practices. Standardized metrics and the insights gained from them will enable the implementation of more effective policies and procedures by the laboratory that are grounded on quantifiable data.

Such metrics can be used to optimize the business processes within a latent print unit to meet their specific requirements. This will allow the laboratory to provide more timely support to its primary customer, the investigators who request identification services, and at the same time help the laboratory to achieve its mission of protecting the public.

### The Way Forward

Following the review of data collection practices of a number of laboratories, we are developing proposed guidelines for standardized latent print casework data collection and metrics. The project seeks to define the core casework metrics a latent print or facial recognition unit should collect in order to gauge effectiveness and efficiency. The purpose is 1) to encourage laboratories (that are not already doing so) to collect and use latent casework metrics, 2) to standardize data collection and metrics so that laboratory managers can have effective ways of comparing work across laboratories, and 3) to encourage LIMS vendors to standardize latent print casework data collection.

The metrics build on the work of FORESIGHT, which is a business-guided self-evaluation of local, regional, state, and federal forensic science laboratories, managed by the WVU College of Business and Economics. FORESIGHT allows laboratory managers to assess resource allocations, efficiencies, and value of services. Our purposes differ somewhat: FORESIGHT has a cost-centric focus, collecting overall laboratory information across all forensic disciplines, by year; we are focusing on collecting a variety of data on latent print operations in much greater detail.

### What Can You Do to Participate?

Success of this project will depend on the willingness of the latent community to embrace this project. This can only be achieved if we can obtain active participation and support from as many laboratories as possible. We would, therefore, like you to be involved: please give us examples of your successes, tell us what metrics are important to you, volunteer to collect pilot data, and review the draft documents as they are produced. If interested, please contact Austin Hicklin ([hicklin@noblis.org](mailto:hicklin@noblis.org)) and Melissa Taylor ([melissa.taylor@nist.gov](mailto:melissa.taylor@nist.gov)).

in justification for increased salaries and increased staffing.

- **Workload Management:** Creating an appropriate staffing model is the primary reason we collect metrics. Once we recognized how many cases come in and how many latents are in each case, and that the amount of AFIS searches have skyrocketed through the years, we were able to create a model that successfully allows us to meet our caseload demands.
- **Effectiveness of processing certain types of data:** Our lab decided to carry out a comparative analysis on the best procedure for cartridges, to determine whether DNA exams or latent print exams were leading to better results. Based on our study, we found out that DNA exams were leading to far more identifications versus latent print exams. Therefore, we decided to only carry out DNA exams on cartridges.
- **Workload management:** One year we had a huge latent print year. As you can imagine this created a significant