Speech Understanding™ Using Deep Learning

Fluency Direct™, M*Modal's top-ranking front-end speech recognition solution, is widely recognized as having the broadest dialectic capabilities and providing superior out-of-the-gate performance, independent of geographic, ethnic or other speech characteristics. To deliver this noteworthy user experience, M*Modal employs a unique, patented approach that combines speech recognition and natural language processing technologies to better understand the meaning, intent and context of the clinician's dictation. This blend of the two technologies is M*Modal's proprietary Speech Understanding which underpins all its speech-driven documentation solutions.

With Fluency Direct, the speech recognition experience of an individual physician is a combination of the collective experience of all M*Modal users and a physician's individually trained speech profile developed over time. M*Modal's approach combines both factorial modeling and deep neural networks (inspired by biological neural networks in animals), which allow the system to represent and model vastly different pronunciations and dictation patterns. This technical capability is paired with actual physician dictations that have been observed over a decade through our cloud-based deployment model. Previous dictations are used to build collective models that reflect the likely things that physicians dictate. While dictating, each utterance is compared to the collective model to validate statements for accuracy based on past documentation. The closer an utterance is to previously known documentation, the more accurate and timely the speech-to-text result returned to the user.

While a user's initial experience is based on the collective experience of all M*Modal users, the system further customizes to the individual user without any active user-driven activity or other training. This is possible because of the system's continuous learning process: the system adapts through observation and continuously improves by capturing the target user's specific pronunciation and dictation patterns. The collected information is combined with data derived from more than 200,000 other physician users. Ultimately, the system automatically builds a highly customized model for each individual user in addition to leveraging a collective model, resulting in game-changing, out-of-the-gate performance.

M*Modal Fluency Direct provides an exemplary user experience with immediate accuracy, speed and performance with multi-layer networks using deep learning technology that:

- Utilizes complex language modelling that is trained with a large data set of existing physician dictations.
- Drives accuracy up and error rates down.
- Accommodates all variety of dialects, individual speaking styles and medical specialties.

To find out more, visit our website at mmodal.com or contact us at 866-542-7253.