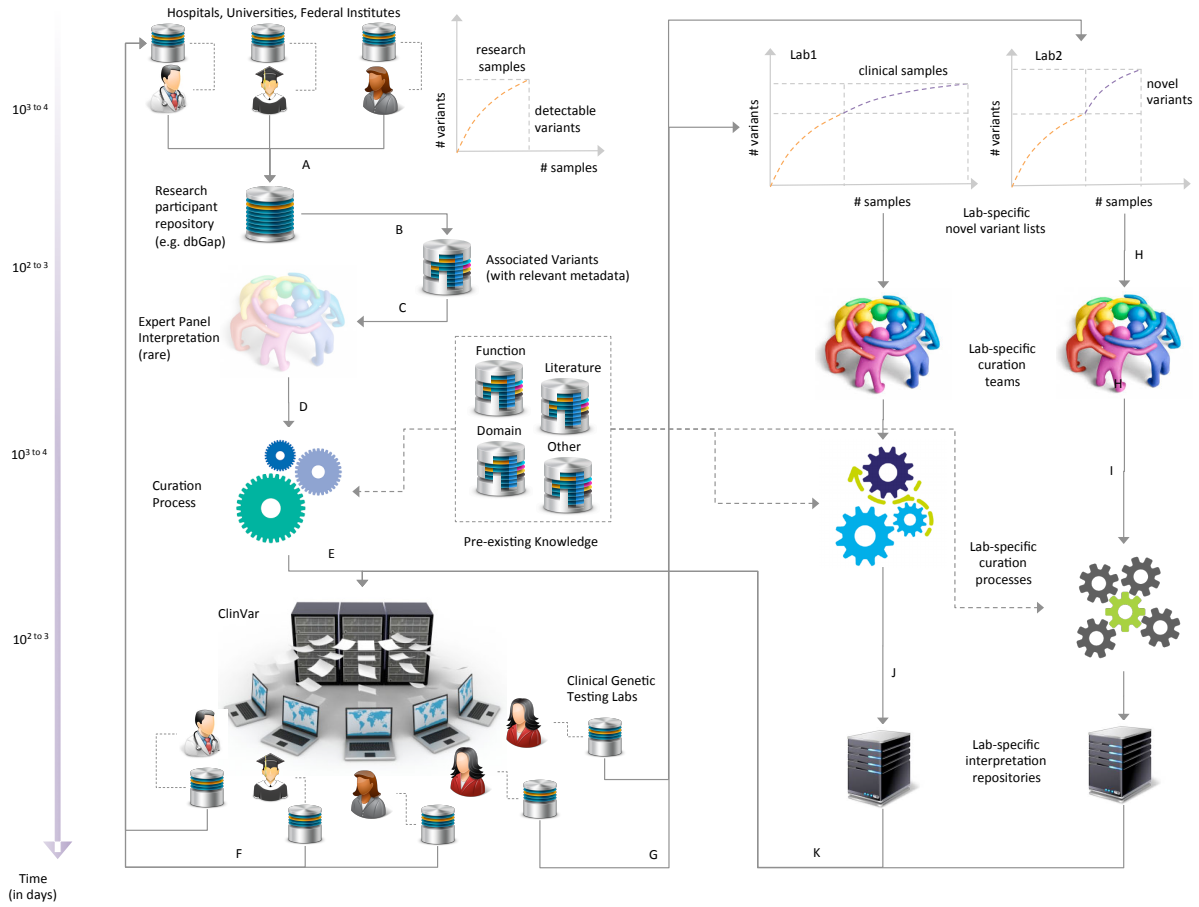

Figure 2. This info-graphic shows the flow of genomic data from research-grade (left) and clinical-grade (right) samples through an interpretation process today, the relevant entities handling such data, and how the results of their interpretation trickle in to a global scientific knowledgebase. **(A)** A research study identifies and enrolls participants at the collaborating institutions' sites, often over a period of years. Participants' individual-level records (genotype data, phenotype data and relevant health data) typically have broad usage consent, and are subsequently aggregated into a central repository. **(B)** Disease-associated variants are revealed through genomic analyses on the dataset, **(C, D)** Rarely, this may be followed by a variant interpretation process that is adjudicated by a panel of experts in the disease domain. More often than not, variants will simply be annotated by automated pipelines before being deposited. **(E)** Variants judged to be clinically significant are then submitted into databanks like ClinVar. **(F)** Any additional samples collected in the interim at the study centers can provide a boost to statistical power, and potentially reveal more disease-associated variants. This necessitates a reiteration through the entire analysis **(A though F)**. Clinical testing laboratories are usually unable to contribute relevant patient samples due to privacy-related safeguards, and are left out of the process thus far. **(G)** Often, novel but "clinically suggestive" variants with no existing clinical interpretations are observed in the patient samples at these testing labs. The amount of suggestive variation observed might differ significantly from lab to lab, due to factors like ethnic representation of patients in the region, number of samples processed by the lab, artifacts due to type of genetic testing technology used, and batch-effects due to undiagnosed errors in sample processing. **(H)** Absent any framework to share patient records or clinical genomic data with the research community, testing labs maintain internal curation teams to perform the task of variant interpretation. **(I)** Curation teams may observe best-practice recommendations to perform triage studies on their novel variants, but pathogenicity assessment can still differ significantly from lab to lab**. (J)** These custom interpretations are subsequently deposited into an internal repository of the testing lab, and might become part of its intellectual property. Unlike the lab's overall curation process, individual interpretations rarely undergo rigorous peer-review. **(K)** Even if these interpretations are subsequently deposited into a community resource like ClinVar, without access to patient phenotype information and a better understanding of the process used to ascertain and annotate these variants at the depositing lab, significant hurdles remain in assessing their true merit.