Promoting Health through Leveraging Clinical and Claims Data

Joshua Baranski

University of Illinois at Chicago

BHIS 593 HI Capstone

Dr. Jacob Krive

December 9, 2022

Promoting Health through Leveraging Clinical and Claims Data

Electronic Data Interchange (EDI) proprietary formats are voluminous among health insurance claims (Office of Assistant Secretary for Planning and Evaluation, 2000). Standardization of health insurance EDI format can help meet the Institutes of Healthcare Improvement's Triple Aim of providing value-based healthcare. How can claims standardization in conjunction with clinical data result in improving population health, reduce healthcare costs and lead to quality and patient satisfaction of care (Institutes of Healthcare Improvement, 2022)?

Through several use cases, we will examine stakeholder involvement, standardization approaches and how to leverage standardization to promote the Triple Aim improvement in providing value-based care. Clinical and claims data can be used to meet the National Committee for Quality Assurance (NCQA's) Healthcare Effectiveness Data and Information Set (HEDIS) measures through a provider dashboard and consumer portal. For example, an illustration of a technological use cases of preventative medicine includes ensuring proper treatment of diabetes mellitus (DMII), post-myocardial infarction treatment and mental health clinical management.

Stakeholder Involvement

Patients

Patients expect to receive value-based healthcare. They want to be satisfied with the care they receive. Furthermore, patients want to be presented with the ability to make informed decisions based upon current health data. They also want to be presented with highest quality of healthcare for their money. As consumers, they do not expect complications to arise from routine services. When complications occur, it results in increased costs and financial strain to the patient. What if patients could be provided risk based referrals based on data from the outcome of providers services from their insurance company? This would provide consumers the ability to make informed decisions on which provider would be most appropriate for their healthcare needs resulting in value-based service. This is only possible when there is standardization among the data for insurance companies to analyze.

Providers

Providers want timely reimbursement for the services they provide. Standardization of claims data can be automated to save time and errors in processing payment reimbursement. The goal is matching the consumer's condition to the specific provider that can provide them the appropriate medical treatment. Standardization of health data to claims data along with analytics can match the right consumer to the most appropriate provider for care. This creates an efficient manner of receiving value-based care.

Payers

Benefits of standardization of claims data along with predictive analytics will provide the consumer with the appropriate physician for treatment. This will provide better outcomes in care. Predictive analytics will lend to decreased costs associated with seeing the right care provider in the most efficient manner. Reducing fragmentation of care enables the consumer to navigate through the healthcare continuum for their healthcare needs. This will lead to decreased costs through preventative healthcare promotion.

Standardization approaches

There are few ongoing standardization use cases that implement HL7 FHIR in conjunction with third party vendor software. There are even fewer use cases that take standardization further by aggregating health data utilizing analytics to improve the efficiency and efficacy of healthcare for consumers (i.e. Oscar).

SmileCDR

This central data repository (CDR) ingests, processes and stores medical data according to FHIR standards. The SmileCDR originally needed a server to run, however it has adopted serverless applications provided by Amazon AWS and Microsoft Azure (Maxi & Morocho, 2022)**.** SmileCDR utilizes extract, transform and load (ETL) to import FHIR data formats into the repository. The data is then stored according to its FHIR format type. SmileCDR states that proprietary formatting is not necessary since data is encoded into FHIR data format types (Rivkin & International Data Corporation [IDC], 2020). However, there are no details on how this is accomplished with unrelated FHIR data formats. Object store and Apache Spark and Hadoop data can be accessed from the repository for analytical measures (Rivkin & IDC, 2020). This access is granted through the FHIR endpoint module, which only grants rights to external users authorized to access the HTTP server integrated with FHIR storage (Rivkin & IDC, 2020). SmileCDR claims to be integrable with various electronic health records (EHRs) (Rivkin & IDC, 2020).

The benefits of SmileCDR is in its integration of structured clinical data. SmileCDR supports FHIR data and integrates the updates into its repository without interruption in data processing (Rivkin & IDC, 2020). The platform upholds private health information (PHI) and personal identifiable information (PII) through assigning specific user rights based on its relationship repository (Rivkin & IDC, 2020). Once rights have been established data can be accessed through Smart FHIR or through an existing integrated repository (Rivkin & IDC, 2020). The scalability potential is another benefit of SmileCDR. The platform can store and convert massive amounts of unstructured data to structured data in FHIR format (Rivkin & IDC, 2020).

There are major disadvantages of SmileCDR. One disadvantage is that SmileCDR only partners with larger healthcare systems (Rivkin & IDC, 2020). It is particularly costly for smaller service providers to embrace. So unless a smaller provider finds a way to partner with a larger service provider this CDR is not feasible. Another shortcoming of SmileCDR is that it does not support marketing, sales or customer service data to capture measures (Rivkin & IDC, 2020). Therefore service providers are then left to third party applications to provide measures to capture this data. This creates fragmentation of data even among agencies within the same company. SmileCDR is trademarked technology that does not provide much information on how it operates. This is so its technological platform can achieve high profitability in the market.

Oscar Health

Oscar Health is unique in its analytical endeavors as it began as an insurance company in the beginning. The goal was to achieve rich analytical data to improve processes of providing high quality care effectively and efficiently to its consumers. Focusing member engagement through incentives can help consumers maintain and improve health. Leveraging Amazon AWS services Oscar Health provided more personized member and provider driven data analytics. This resulted in reducing costs while promoting consumer healthcare outcomes. The consumer decisions were enabled through the aggregation of data among providers and treatment outcome analysis. The data was presented in an easy to understand chart format to consumers that illustrates how to take appropriate measures toward a healthier state.

Oscar Health utilizes Amazon AWS for various applications however not all of their platform is disclosed due to market secrecy to retain financial support. Oscar Health uses Amazon Elastic Compute an open cloud based system that runs CentOs operating system (Barr, 2015). This houses Amazon Machine Image (AMI) which self-configures while booting up

Ansible (Barr, 2015). Ansible is an open source automating software utilizing python that leverages workflows, app deployment and system uploads (Barr, 2015). This software is the backbone of Oscar Health platform. This enabled Oscar to drive analytical insights in a more effective and efficient manner. The integration of Consul server discovers the appropriate data clusters to network across cloud networking (Barr, 2015). Pairing Amazon Route 53 the network traffic is able to handle the enrollment traffic surges (Barr, 2015). The Oscar Health platform ingests, cleans, and uploads data on availability of providers in the insurance network paired with member health needs (Molteni, 2017). This allows members the ability to make decisions based on the right treatment from the right provider contributing to efficacious value-based care. Oscar is more than just a claims payment system, it utilizes analytical data in real-time to drive healthcare insights (Thompson, 2018). Real-time insights enable the consumer to transition across the healthcare continuum without fragmentation in care resulting in poor health outcomes.

Promoting Triple Aim

To promote the Triple Aim in the health insurance sector there needs to be standardization among disparate silos of health data. Once standardization is achieved then, health data can be examined and algorithms designed for the promotion of value-based consumer care. It has been examined how Oscar has provided its consumer base with value-based care utilizing its data repository for predictive data analytics of provider treatment claims (Molteni, 2017). Promoting the Triple Aim provides the basis in the construction of the preceding architectural diagram. This architectural diagram will serve as a foundation that health insurance companies can use to leverage standardized data like Oscar to promote value-based care in an efficient and efficacious manner.

Architectural Design

The architectural design for standardization of claims processing in health insurance should be composed of multiple components. The utilization of HL7 FHIR server is necessary to align claims data with health care data from the provider allowing easy codification and analysis. The architecture is composed of data extraction, transform and load components (ETL) from the health insurance business data feed as well as the HL7/FHIR healthcare data feed.

Health Insurance Business ETL

Business data such as claims forms, insurance identification and invoices are sent to Amazon AWS cloud service to the Amazon S3 and arranges different collected documents in the organization (Condello et al., 2022). Next, AWS Glue orchestrates exchange, transform and load (ETL) operations as data becomes available from Amazon S3 allowing the data to be queried. Then, AWS Fargate Edifecs X-Engine takes the documentation from AWS Glue and validates and classifies the various financial and business forms (Condello et al., 2022).

HL7/FHIR Healthcare ETL

HL7/FHIR healthcare data is transferred from EHRs utilizing coding imported to a HL7/FHIR server. Amazon S3 triggers data processing of the health data. Next, AWS Glue (ETL) makes data easy to discover, combine and prepare for predictive analytics, machine learning (ML) and application development (Condello et al., 2022). AWS Lambda then extracts details on whether data is structured or unstructured. The unstructured data is sent to the AWS Step Function that extracts the unstructured data to create meaningful use applications (Mallick, 2019).

Documentation Classification

Amazon Textract is a ML application that reads and processes documents extracting text, handwriting, tables with automation for further classification. Amazon Comprehend a natural language processing (NLP) application that gains data insights from text (i.e. person, location, data, etc.) (Condello et al., 2022). It also detects the language, personally identifiable information (PII), personal health information (PHI) and allocates the data into relevant classes with confidence scores (Condello et al., 2022).

Documentation Extraction

Amazon Textract at this step is utilized to process data efficiently and effectively to gain insights from the data (e.g. claims forms, medical forms, etc.). Comprehend then extracts this data and transcribes to a JSON format to create document formats that are universal across various application processing interfaces (API) (Condello et al., 2022).

Document Enrichment

The main goal of Comprehend, is to redact PII and PHI to conform to the Health Insurance Portability Accountability Act (HIPAA). This ensures compliance prior to the data access after analytical insights are obtained (Narayanan et al., 2022). Comprehend Medical is a NLP conformed to HIPAA standards that utilizes ML to extract medical data (e.g. medical text, prescriptions, diagnosis etc.) and medical ontologies (e.g. ICD-10CM, RxNorm, CPT, SNOMED CT codes) (Narayanan et al., 2022).

Analytics Workflow

After enrichment, data is transferred to Amazon Redshift. Amazon Redshift is a data warehouse solution in the cloud that stores data prior to it undergoing analytical measures (Narayanan et al., 2022). Amazon Kinesis Analytics takes the stored data from the warehouse and aggregates meaningful use of the data utilizing predictive and prescriptive analytics (Narayanan et al., 2022). This would help identify relational trends in the data.

Data Access

After analytical measures of abstracting meaningful use of the data, it is then stored in the Amazon RDS database. The Amazon RDS database has an API that allows consumers to interact with the information on a SPA website.

Architectural Advantages

Constructing a ML predictive analytics model has the ability to leverage value-based services based on providers' practice. This will include the treatment provided by the provider not necessarily their job title, which can be misleading. Also, infection rates can be analyzed alongside the provider's treatment. Comparing providers' evidence-based treatment outcomes data via a digital scorecard to consumers can allow them to make better informed value-based care decisions.

The architecture lends itself to the ability to standardize data across platforms as it supports and transcribes various data types. This is useful in identifying business documents (from EDI) that may be in other formats to standardize health data and store in the data warehouse. Formats can be quickly mapped to a standardized format type. The standardization and analysis of data from claim forms, as well as the ontological medical coding formats, can help to prevent fraudulent claims that increase healthcare costs.

The Amazon AWS architecture not only lends the ability to standardize data, but it is also versatile in its ability to scale to the specific insurance industry's size and needs. For value-based healthcare services, the use of the cloud reduces costs associated with maintaining on-premise hardware and software, as well as costs associated with staffing personnel to troubleshoot any issues that arise. Health insurance companies have a vast array of data to store and analyze. On-premise technology has the potential to become expensive in upscaling to address healthcare needs.

Applications

Through the utilization of prescriptive analytics, clinical and claims data identifies the HEDIS measures of newly diagnosed diabetic (DM2) patients (ICD-10 E11). Hemoglobin A1c (HbA1c) clinical results examine if the consumers' diabetes is controlled (<8%) (e.g. HCPCS G8016) (National Committee for Quality Assurance, n.d.-b). Following HEDIS measures, data can determine if the consumer has had a scheduled eye exam performed (e.g. ICD-10 E10.39, E11.39), blood pressure (e.g. HCPCS G8024) has been addressed or if medical treatment has been obtained for nephropathy (e.g. ICD-10 E10.21, E11.21) (National Committee for Quality Assurance, n.d.-b). Consumer usage of the portal, via a case manager ensures the consumers have scheduled appointments along with access to their diabetic supplies (e.g. glucometer, lancets: HCPCS A4259). Analyzing clinical and claims data allows the insurance company to focus on consumers that haven't received the necessary treatments (e.g. RxNorm 151827), procedures or medical equipment. Follow-up appointments with nutritionists (e.g. CPT 97802, 97803, S9470) and provider appointments (e.g. CPT 99204, 99396) can also be confirmed for health promotion. Patients who have been identified as pre-diabetic (e.g. ICD-10 R73.09) can access the portal for referrals to a nutritionist to help avoid potential future diabetic related issues. The provider dashboard allows physicians to address any lapse in treatment and provide the consumer with education for health promotion. The technological architecture can effectively and efficiently address comprehensive diabetes care and curbing the occurrence of serious complications in real-time.

Prescriptive analytics also addresses hospitalized patients with a diagnosis of acute myocardial infarction (MI) (e.g. ICD-10 I21). Beta-blockers (e.g. RxNorm 203344, 224909) are used in the treatment and prevention of acute MI and eventual heart disease. Utilizing clinical and claims data allows the tracking of beta-blocker medication of refills for six months post-discharge, which is the preferred preventative treatment for MIs (National Committee for Quality Assurance, n.d.-d). Leveraging clinical and claims data can also check for repeat procedures such as ECGs (e.g. CPT 93229, 93010, 93005). This data can be accessed through the patient portal for health care consumer reminders. This could include a referral to cardiology or follow-up appointment with main provider (e.g. CPT G0406, G0407, G0408, G9968; ICD Z71). For health consumers who are unable to utilize the portal, the data is also routed to the case manager or provider who prescribed the medication for real-time education and follow-up.

This architecture can also address the use case of depression follow-up after emergency department visit as well as medication management. The NCQA HEDIS measure for health consumers' with a mental health/depression (e.g. ICD-10 F32, HCPCS G0444) diagnosis have a follow-up provider appointment within a week and another within 30 days after discharge (e.g. CPT 90791, 90792) (National Committee for Quality Assurance, n.d.-c). Clinical guidelines recommend medication adherence for 12 weeks for acute episodes and six months for continuation of treatment (National Committee for Quality Assurance, n.d.-a). Through prescriptive analytics the clinical and claims data is leveraged to promote antidepressant medication compliance (e.g. RxNorm 58827) through monitoring effectiveness of treatment with the minimalization of side effects (e.g. ICD-10 T43) as displayed through the clinical dashboard and consumer portal.

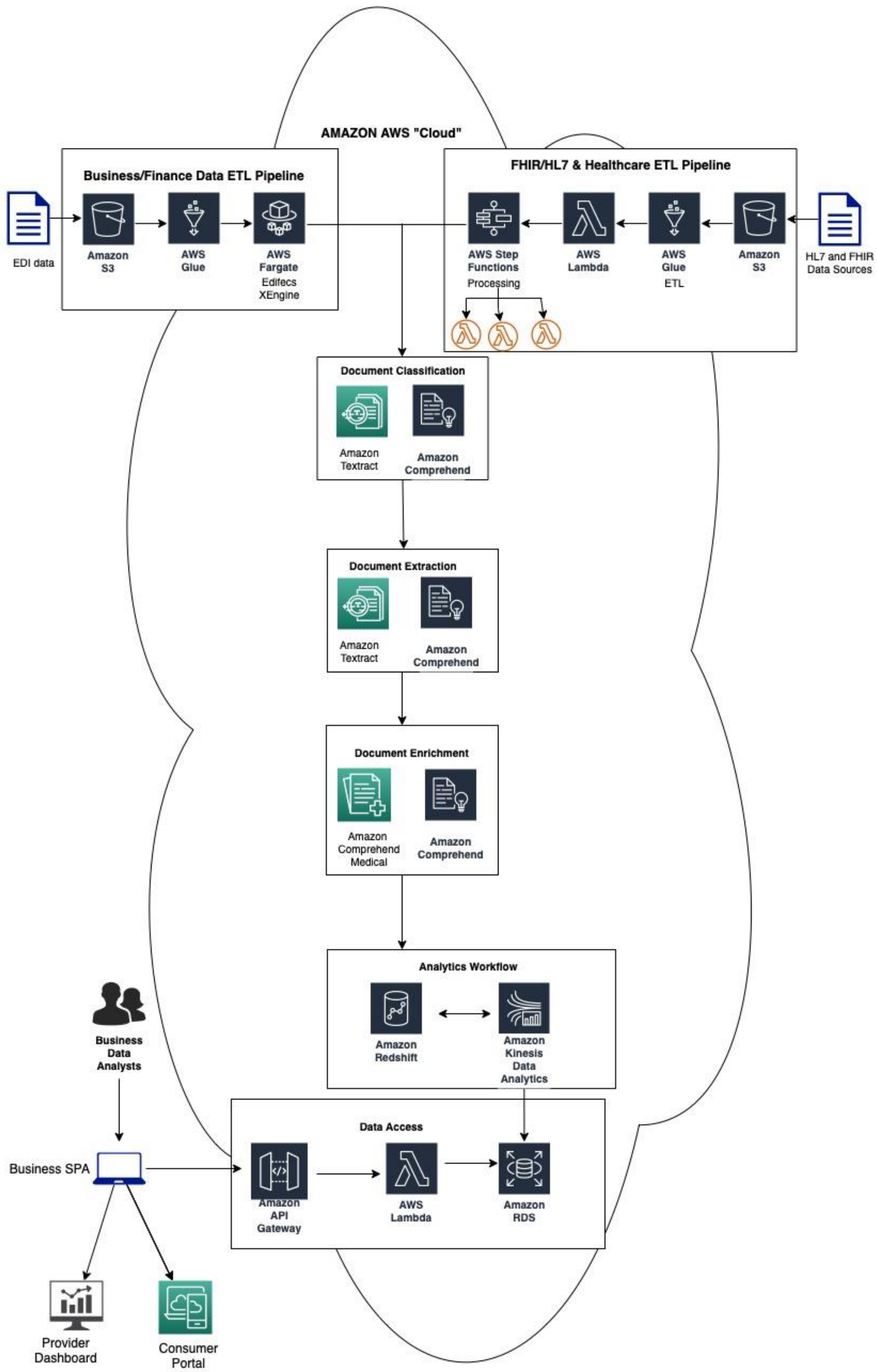
Architectural Outcomes

Consumer clinical and claims data merges in through the ETL pipeline and undergoes transformation through several processes. Aggregation of structured/unstructured of consumer clinical and claims data happens through data classification, extraction and enrichment phases.

Also during these phases, the data is assigned confidence scores as well as being transcribed to JSON for interoperability. The aggregated consumer clinical and claims data is stored into Amazon Redshift Data Warehouse prior to analytical processing. Amazon Kinesis performs analytical processing deriving insights on the clinical and claims data. It is during this process where missing or unrepresented data based on the aforementioned HEDIS measures are sent to the provider dashboard and consumer portal. This allows for education and opportunity to bridge fragmentation in care. At the SPA the data is automated for real to near-real time purposes to provider dashboard and patient portal. At this point manual review by data analysts can also be performed for data confidence purposes. Data at the provider dashboard and consumer portal enables real-time or near real-time education addressing gaps in HEDIS treatment measures. This includes any missing lab work, device purchases, provider appointments, medication (refills) and diagnostics resulting in fragmentation of the patient care continuum. The dashboard and consumer portal also aligns the consumer with the appropriate provider based on diagnostic codes.

Conclusion

The designed architecture leverages clinical and claims data to address the continuity of care in an effective and efficient manner. The dashboard information promotes improved health and wellness through addressing preventative and continuation of care measures and reducing the overall burden of lapses in treatment resulting in 30-day readmissions. The closure of lapses in treatment through providing education measures by addressing overall population risk factors, will decrease overall healthcare costs. Through this architecture several HEDIS use cases have been generated and shown how effectiveness and efficiency can contribute to value-based care.



AMAZON AWS "Cloud"
Business/Finance Data ETL Pipeline
EDI data
Amazon S3
AWS Glue
AWS Fargate
Edifecs XEngine
FHIR/HL7 & Healthcare ETL Pipeline
AWS Step Functions
Processing
AWS Lambda
AWS Glue
ETL
Amazon S3
HL7 and FHIR Data Sources
Document Classification
Amazon Textract
Amazon Comprehend
Document Extraction
Amazon Textract
Amazon Comprehend
Document Enrichment
Amazon Comprehend Medical
Amazon Comprehend
Analytics Workflow
Amazon Redshift
Amazon Kinesis Data Analytics
Business Data Analysts
Business SPA
Data Access
Amazon API Gateway
AWS Lambda
Amazon RDS
Provider Dashboard
Consumer Portal

References

Barr, J. (2015, May 5). *Oscar Health – a new kind of health insurance company, powered by aws*. Amazon Web Services. https://aws.amazon.com/blogs/aws/oscar-health-a-new-kind-of-health-insurance-company-powered-by-aws/

Condello, T., Rane, C., & Sahu, S. (2022, November 3). *Intelligent document processing with AWS AI services in the insurance industry: Part 1*. Amazon Web Services. https://aws.amazon.com/blogs/machine-learning/part-1-intelligent-document-processing-with-aws-ai-services-in-the-insurance-industry/

Institutes of Healthcare Improvement. (2022, October 26). *The IHI triple aim*. https://www.ihi.org/Engage/Initiatives/TripleAim/Pages/default.aspx

Mallick, M. (2019, October 3). *Achieve healthcare interoperability by integrating Amazon comprehend medical with FHIR*. Amazon Web Services. https://aws.amazon.com/blogs/industries/achieve-healthcare-interoperability-by-integrating-amazon-comprehend-medical-with-fhir/

Maxi, K., & Morocho, V. (2022). Integrating medical information software using health level seven and FHIR: A case study. In *Smart technologies, systems and applications* (pp. 84–98). Springer International Publishing. https://doi.org/10.1007/978-3-030-99170-8_7

Molteni, M. (2017, May 12). *How one startup built better health insurance with the magic of data*. WIRED. https://www.wired.com/2017/05/one-startup-built-better-health-insurance-magic-data/

Narayanan, U., Rane, C., & Sahu, S. (2022, November 3). *Intelligent document processing with aws ai and analytics services in the insurance industry: Part 2*. Amazon Web Services.

https://aws.amazon.com/blogs/machine-learning/part-2-intelligent-document-processing-with-aws-ai-and-analytics-services-in-the-insurance-industry/

National Committee for Quality Assurance. (n.d.-a). *Antidepressant medication management - NCQA*. NCQA. Retrieved November 17, 2022, from https://www.ncqa.org/hedis/measures/antidepressant-medication-management/

National Committee for Quality Assurance. (n.d.-b). *Comprehensive diabetes care - NCQA*. NCQA. Retrieved November 17, 2022, from https://www.ncqa.org/hedis/measures/comprehensive-diabetes-care/

National Committee for Quality Assurance. (n.d.-c). *Follow-up after hospitalization for mental illness - NCQA*. NCQA. Retrieved November 17, 2022, from https://www.ncqa.org/hedis/measures/follow-up-after-hospitalization-for-mental-illness/

National Committee for Quality Assurance. (n.d.-d). *Persistence of beta-blocker treatment after a heart attack -NCQA*. NCQA. Retrieved November 17, 2022, from https://www.ncqa.org/hedis/measures/persistence-of-beta-blocker-treatment-after-a-heart-attack/

Office of Assistant Secretary for Planning and Evaluation. (2000, October 15). *Health insurance reform: Standards for electronic transactions*. ASPE. https://aspe.hhs.gov/reports/health-insurance-reform-standards-electronic-transactions

Rivkin, J., & International Data Corporation. (2020, December). *IDC Marketscape: U.S. customer-360 data platforms for payers 2020–2021 vendor assessment* (#US46997020e) [Report]. IDC: The premier global market intelligence company. https://www.idc.com/getdoc.jsp?containerId=US46997020

Thompson, N. (2018, August 14). *Health care is broken. oscar health thinks tech can fix it*. WIRED. https://www.wired.com/story/oscar-health-ceo-mario-schlosser-interview/