# **Metadati incontro con azienda**

**Data incontro:** 20/10/2023

**Partecipanti:** All (Tranne Riccardo Alberto Costantin)

**Partecipanti azienda:** A.Dorigo, F.Pallaro, D.Zorzi

**Tempo:** 30 min

# **Email di contatto**

To: d.zorzi@synclab.it andrea.dorigo@synclab.it Cc: f.pallaro@synclab.it

Oggetto: Richiesta incontro Software Engineering UNIPD [SyncCity]

Gentili referenti di SyncLab,

siamo il gruppo SWAT Engineering, studenti del corso di Ingegneria del Software, dell’Università degli Studi di Padova, partecipanti al primo lotto dell’anno 2023/24.

Con la presente email vogliamo esprimere il nostro interesse riguardo al Vostro capitolato e saremmo molto interessati ad approfondire il progetto InnovaCity, vorremmo pertanto chiedere un incontro conoscitivo al fine di chiarire alcuni dubbi.

Siamo flessibili per quanto riguarda la data e l'orario dell'incontro e siamo disposti a conformarci alle Vostre disponibilità.

Ringraziamo anticipatamente per l’attenzione, attendiamo Vostre notizie.

Cordiali saluti

# **Domande & Risposte relative**

1. Se e' possibile usare linguaggi diversi rispetto a Python per i simulatori (sperimentazione e usare linguaggi resource efficient per l'IOT).  
     
    Si può utilizzare qualsiasi linguaggio si voglia (la sperimentazione è la benvenuta). Python è stato consigliato semplicemente perché, pur non essendo il linguaggio più efficiente, lo è effettivamente nella scrittura e consente di generare centinaia di dati al secondo in maniera molto semplice. L’attenzione deve essere comunque posta sul realismo dei dati simulati: ad esempio, volendo simulare la temperatura esterna di un edificio, i dati dovrebbero avere un andamento sinusoidale in base all’orario del giorno. In questo senso, se Python offre molteplici librerie che potrebbero risultare utili a tale scopo, linguaggi più recenti come Rust potrebbero non offrire librerie altrettanto specifiche. Volendo, si possono sfruttare anche API che forniscono grossi quantitativi di dati. L’obiettivo rimane quello di dimostrare la capacità di ottenere e gestire una mole di dati consistente o, in altre parole, la componente Edge più vicina ai sensori IoT.
2. Quando parlate di formazione deep-dive che cosa si intende? È basato su una specifica tecnologia dello stack?  
     
    La formazione è aperta a qualsiasi richiesta proveniente dal gruppo, in forma di sessioni Q&A e bootcamp su argomenti a piacere. Le sezioni dell'architettura ritenute più complesse dal proponente sono Kafka, in quanto basato su event-based architecture, e ClickHouse, in quanto e’ un database colonnare che funziona tramite segmentazione. La proponente si offre di fornire delucidazioni anche sulla gestione della containerizzazione: in particolare, si suggerisce di inserire il tutto all’interno di un docker composed, anche per poter implementare un event-based architecture riproducibile con un singolo click (indubbiamente un valore aggiunto a livello di dimostrazione del prodotto).
3. Quando si parla di big data, che mole di dati hanno in mente.  
     
    Non si parla di vero e proprio big data (decine di migliaia di dati al secondo da gestire in modalità distribuita/parallela), nonostante le macchine moderne siano in grado di generare ed elaborare centinaia di dati al secondo senza problemi. Aggiungere un problema di scalabilità non sarebbe necessariamente utile al fine di dimostrare le competenze tecniche del gruppo nel costruire l'architettura proposta. L’ordine di grandezza dei dati si aggirerà dunque intorno ai 100 dati/secondo. Naturalmente, dipende anche dal tipo di sensore simulato: ad esempio, ci si aspetta che un sensore di temperatura esegua all’incirca 1 rilievo/minuto.
4. Chi è/sono gli utenti finali dell'applicazione e capire se ci sono dati "sensibili".  
     
    Dal punto di vista reale, se si trattasse di sistemi di monitoraggio e business intelligence di grandi aziende l’applicazione verrebbe utilizzata dalla sezione rilevante dell’azienda stessa; se si trattasse invece di un progetto pubblico, questa verrebbe probabilmente utilizzata dalla municipalità interessata. Volendo delineare un target su cui impostare certe logiche di user experience, la proponente suggerisce di immaginare che l’utente principale sia costituito da tecnici del comune di Padova che hanno necessità di monitorare e riportare la situazione della città ad altri enti. Ad esempio, per dei geometri puo’ essere utile fare un’analisi della viabilita’ delle biciclette elettriche di Padova (servizio offerto dal comune, a pagamento).
5. Qual è la parte piu complessa del progetto secondo loro.  
     
    La parte più complessa da comprendere è sicuramente quella relative a ClickHouse (query-engine database di tipo OLAP). Inizialmente, tuttavia, la difficoltà potrebbe essere legata all’individuazione del tipo di dati da generare (inventarsi sorgenti di dati sensate nel contesto proposto e tentare di simularne il comportamento). Occorre trovare un compromesso tra simulare una base di dati che sia logicamente utilizzabile (e dimostri appieno le capacità del prodotto) e ricreare un contesto familiare per elicitare un feedback positivo da parte del cliente. L’insieme minimo di sensori da simulare è comunque presente all’interno del capitolato, ma eventuali aggiunte sono più che benvenute. La proponente si offre anche di fornire esempi che illustrino il modo in cui solitamente è strutturato il file JSON contenente gli input dei sensori (una volta mascherati alcuni dati in ingresso). Viene inoltre chiarito che i requisiti di prodotto non includono componenti di data analysis che evidenzino correlazioni tra dati di natura diversa (es. situazione metereologica - utilizzo di biciclette). Tuttavia, l'aggiunta di tale funzionalità sarebbe molto apprezzata. In tal caso, si consiglia di impleInolmentarla mediante Grafana anziché utilizzare software come Apache Superset, che potrebbero risultare più complessi.
6. Quanta parte di data analysis c'è nel progetto.  
     
    sLa componente di data analytics non è stata inserita tra i requisiti obbligatori nel capitolato onde evitare un layer di complessità aggiuntivo, ma si tratta esattamente di ciò che verrebbe fatto per finalizzare il modello. Grafana è uno degli strumenti ideali per la visualizzazione dei dati, in cui l’analisi si potrebbe fare comparando query diverse e mostrando correlazioni interessanti (ad esempio, visualizzare le linee di correlazione tra i dati di utilizzo delle biciclette elettriche ed eventi metereologici). Volendo, si possono utilizzare altri strumenti quali SuperSet o PowerBI per fare data visualization, in quanto sono in grado di gestire dati generati da meccaniche di machine learning collegate a moduli esterni in modo da trovare eventuali correlazioni tra i dati.
7. Quanta disponibilità danno nel caso si chiedesse la presenza nelle Sprint Review.  
     
    La proponente prevede di utilizzare un calendario condiviso su Google in modo da comunicare anticipatamente le date degli incontri SAL (da effettuare sia da remoto che in sede), per ognuno dei quali verrà messo a disposizione un file condiviso contenente l’ordine del giorno. Per quanto riguarda la partecipazione agli Sprint Review (in presenza del committente), la proponente si offre di mettere a disposizione un contatto che partecipi in presenza qualora fosse necessario. Di solito viene impiegato Discord come piattaforma di comunicazione principale, perciò ogni anno vengono creati dei canali appositi per la comunicazione tra proponente e gruppo, ma anche tra gruppi distinti qualora si trovassero a lavorare sullo stesso capitolato. Inoltre, si suggerisce di anticipare le eventuali domande da porre prima di ogni incontro via email, in modo da ottenere un riscontro quanto più completo e tempestivo. La proponente è disponibile per eventuali incontri in sede, previa richiesta anticipata, con l’obbiettivo di favorire una migliore collaborazione con il team.
8. Quale disponibilità di hardware e software(servizi) mettono a disposizione (stress test).  
     
    La proponente sta ancora discutendo interamente di quali strumenti mettere a disposizione del gruppo: l’obiettivo è quello di offrire un ambiente per effettuare installazioni da remoto, soprattutto durante la fase finale di sviluppo. Per quanto riguarda i servizi software, la proponente suggerisce di creare un progetto su GitHub in modo che sia accessibile ad ambo le parti con la possibilità di renderlo completamente pubblico una volta ultimato a discrezione del gruppo (viene specificato che la proprietà del codice è del gruppo). Jira viene sconsigliato nonostante offra molte più funzionalità rispetto a GitHub poiché, data la complessità nell’apprenderne l’utilizzo, potrebbe dilungare i tempi di sviluppo. Come project management system viene consigliato ClickUp in quanto richiede un breve periodo di apprendimento e permette di mantenere il focus sullo sviluppo del progetto. Altre alternative includono Notion o Trello.
9. (Domanda per chiarire) Che nome dobbiamo usare sui documenti: SyncLab o InnovaCity?  
     
    InnovaCity.