

# Ülikoolist andmeteadusse - kuidas alustada?

Mikhail Iljin

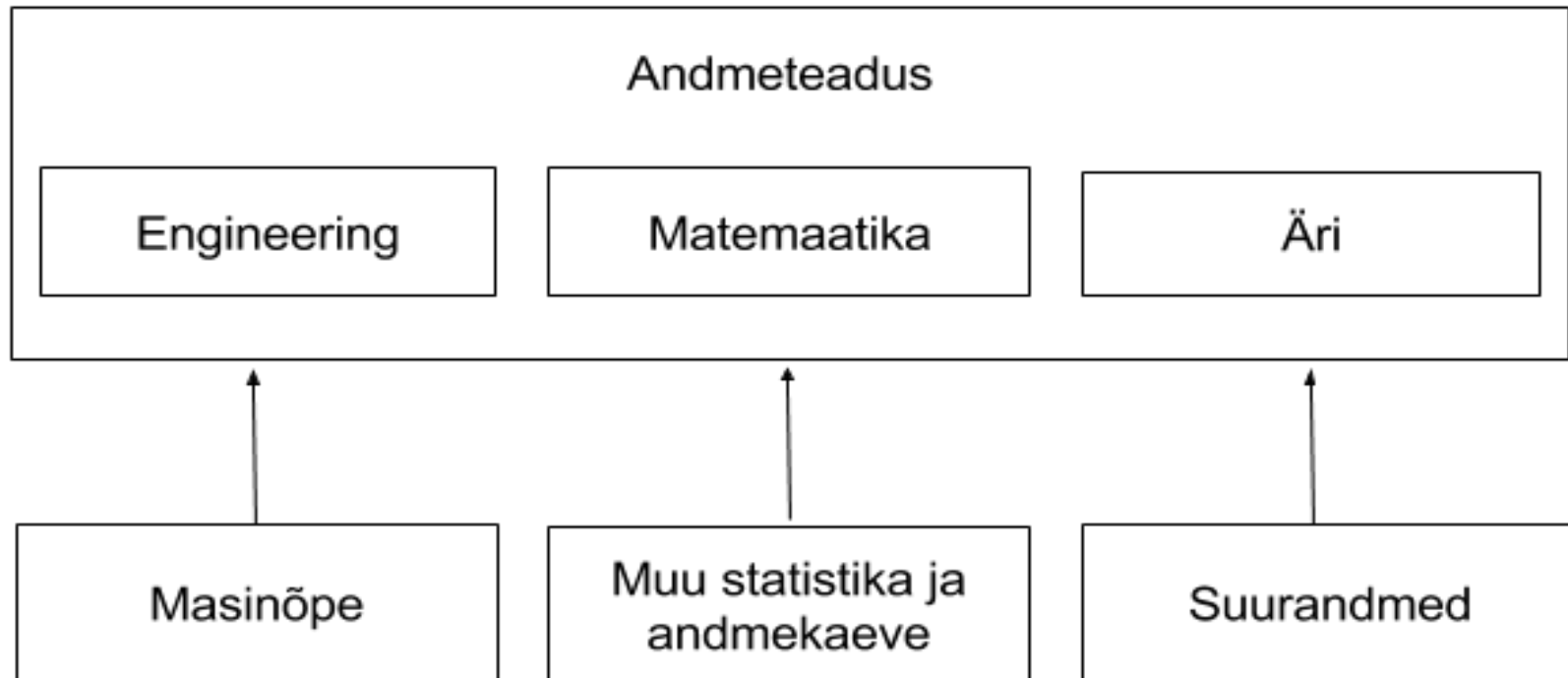
# Kes ma olen?

- Trojaanid
- PHP saidid
- Veebiportaalid Javas, Swedbank internetipank
- SK: ID-tarkvara, Mobiil-ID, digidokumendid
  
- Nortal big data/data science tiim
- FeelingStream startup
- Mooncascade big data/data science tiim
- Mõned freelance projektid siin ja seal

# Terminid ja buzzwordid

- Artificial intelligence
  - Ei ole veel sellist asja
  - Teooriat ei ole, aga midagi proovitakse ja tehakse imesid
    - Kas jätkusuutlik?
- Data science
  - 1/3 engineering
  - 1/3 matemaatika
  - 1/3 äri
- Big data
  - Mis ei mahu mällu ega relatsioonilisse andmebaasi
  - Ebestandartsed formaadid
- Machine learning

# Millest töö koosneb? Näited?



# Millist haridust/oskusi on vaja?

- Big data – puhas inseneritöö
- Masinõpe – inseneritöö + osaliselt matemaatika
- Statistika ja “päris” andmeteadus – väga palju matemaatikat (alustamiseks piisab baka tasemel)
- “Applying data scientist” - 90%
  - Tavaliselt magister
- “Researching data scientist” - 10%
  - Tavaliselt doktorant/doktor
- “Data engineer”

# Oled IT-tudeng või alustav arendaja

Oletan, et sa:

- **Engineering:** Oskad natuke programmeerida Pythonis ja Javas
- **Engineering:** Tead mis on relatsiooniline andmebaas (MySQL, PostgreSQL, Oracle jt)
- **Matemaatika:** Tegid ära Matanalüüs I ja II
- **Äri:** Tegid mõni projekte tiimis (ja tead, kui keeruline võib inimestega olla)

## **Sarnasus**

- Programmeerimine, aga kood ei ole lõpptoode

## **Erinevused**

- Väga väga palju suhtlust
- Kiire prototüüpimine – liiga hea on halb

## **Lisaks**

- Funktsionaalsed keeled (R, Python, mõnikord Scala)
- NoSQL andmebaasid
- Hajutatud süsteemid
- Matemaatika. Kust alustada?
- Teaduspaberid

# Kes tegeleb?

- Ettevõtted

- Mooncascade, Nortal, Proekspert, STACC
- Starship, TransferWise, Alphablues, Pipedrive
- Ja paljud teised: <http://datasci.ee/ressursid/>

- Ülikoolid

- Tartu Ülikool – biotech, brain-computer interface, mobility labs, ...
- Tallinna Tehnikaülikool – speech tech lab, ...



# mooncascade.com/career



## IOS SOFTWARE ENGINEER

---

You know your way around Xcode, Swift, and the iOS SDK. You have the newest iPhone with the latest iOS installed. If this sounds like you, then we obviously have lots of interesting stuff to discuss. Apply and let's have a chat in our Tallinn or Tartu offices!



## ANDROID SOFTWARE ENGINEER

---

Do you speak Java? Do you regularly negotiate deals with Activities and Services? Do you become alerted when you hear someone say Kotlin on the street? If your answer to these questions is "yes", then we obviously must meet with you! Apply before it's too late!



## WEB SOFTWARE ENGINEER

---

You know which is the best PHP framework that you have used. You can tell between HTML5 and HTML. The last time you used the SQL EXPLAIN command was yesterday. If the above brings about positive emotions, then we have lots of common topics to cover. Apply and let's meet!



## SENIOR PROJECT MANAGER

---

We're looking for a brilliant juggler who can keep their calm and manage projects in spite of even the most unexpected challenges. If catching a tiger running loose or standing an eye to eye with a raging bull doesn't scare the hell out of you, you're ready to be challenged. Apply and let's get this show started!



## SENIOR UX/UI DESIGNER

---

The best design is the simplest one that works or better yet, has no UI at all. If you want to create magic and surprise users with an awesome experience, and you have an eye for a clean, top-notch visual design, you might just be the one we're looking for! Copy the link of your impressive kick-ass online portfolio to your clipboard and apply now!



## SENIOR QA

---

Others say you break stuff. You say it was already broken when you got it. You know that Selenium is not only a chemical element and that there is Michael "Not That Singer" Bolton. And you're passionate about improving quality and building the best possible product. Yes? Then apply and let's get acquainted!

[https://www.cvkeskus.ee/view\\_jobad.php?job\\_id=441490&w=1](https://www.cvkeskus.ee/view_jobad.php?job_id=441490&w=1)

### **Ettevõtte pakub**

- Brutopalka 4000 EUR kuus
- Arenguvõimalust suurandmete analüütika vallas
- Olulist tööd heas tiimis, mis aitab äridel kasvada
- Suurepärasest töökeskkonda
- Võimalust töötada kiiresti arenevas ja moodsas firmas

<https://www.stacc.ee/vacant-position-senior-data-scientist/>

### **What we offer you:**

- Salary of €3000+, depending on experience and competence;
- Highly qualified colleagues with necessary know-how;
- Small development teams and a dynamic work environment;
- Flexitime and mobile work;
- Participation in training courses within your area of work;
- *Cozy office at the city center of Tartu*

Välismaal kordades suuremad

- **Miinused**

- Väike seltskond, kõik tunnevad kõiki - palju töökohti pakutakse tuttavatele
- Kõik tahaksid tegeleda närvivõrkudega – äris vähe vajadust, rohkem akadeemias
- “To apply” - 90%, “to research” - 10%
- Klientidel on väiksed eelarved, tihti kukkub ära

- **Plussid**

- “Rocket scientist” tunne
- Väike seltskond, kõik tunnevad kõiki – sind märgatakse väga kiiresti
- Valdkond lapsepõlves – palju mänguruumi
- Väärtuslik äri- ja läbirääkimiste kogemus

Kuidas sinna sattuda?

Millest alustada?

- Haridus, haridus, haridus! – aga päris kaugel
- Või tõsine töökogemus – aga vajab haridust
- Või et sa tunneksid inimesi, sind märgataks ja selle ajani sul ei kaoks huvi ära

# Kodune klassika

- Machine Learning (4 kuud, 8 tundi nädalas)  
<https://www.coursera.org/learn/machine-learning>
- Data Science Specialization (9 kuud, 4 tundi nädalas)  
<https://www.coursera.org/specializations/jhu-data-science>
- Masinõppe võistlused  
<https://www.kaggle.com/>
- Masinõppe uudised ja arutelid  
<https://www.reddit.com/r/datascience/>
- Kodused projektid, Github konto

Mis sellega vaalesti on?

<http://datasci.ee/>

<https://tallinn.city.ai/>

<https://www.meetup.com/Machine-Learning-Estonia/>

# Kaval karjäär

Kandideeri inseneriks sinna, kus tehakse andmeteadust – 1 hiljuti, 1 kunagi, 1 mina



# Motivatsioon

- Github häkid – blogisse
- Häkatonid
  - <https://www.facebook.com/events/169157423649172/>  
Garage48 SpaceTech: Integrated Applications, 3-5. november
  - <https://www.facebook.com/events/356918158098210/>  
Social Impact Data Hack 2017, 10-12. november
  - <https://hackjunction.com/>  
Junction 2017, 24-26. november

# Motivatsioon

- Startupid
  - Kasvavad häckatonidest välja
  - Viivitamatult saad andmeteadlaseks – ja saad reaalselt kogemust
  - Minu kogemus: kaks hackatoni osalejana, 1.5 startupi – kuid vajab pühenduvust
  - Startupide kiirendid aitavad sind
  - Kõik (ka investorid) praegu armastavad n.ö. AI projekte

# Motivatsioon

- Atmosfäär
- Inimesed
- Asi meeldib + On kellega tegeleda + Tegeled = Kogemus
- Kogemus → Varsti ka raha, IT-s eriti

Kõik

Meilisari:

<http://datasci.ee/sissejuhatus/#scrollto-signup-form>

















