



**日本音響学会2024年秋季研究発表会
ビギナーズセミナー資料
「音響研究者のキャリアパスについて」**

2024年9月4日

三村 正人

日本電信電話株式会社 人間情報研究所

職歴



■ 前職 京都大学

- ◆ 2006.4～2014.7 京都大学 学術メディアセンター 研究員
- ◆ 2014.7～2023.6 京都大学 情報学研究科 研究員
 - › 2021年9月 京都大学から博士（情報学）の学位取得

■ 現職 日本電信電話株式会社

- ◆ 2023.7～ 日本電信電話株式会社 人間情報研究所 研究員

■ 参考：音声業界の友人たちのよく見るキャリア

- ◆ 修士で日本企業の研究職に就職
- ◆ 社会人博士として大学で学位取得
 - › 受け入れ先探し、仕事との両立、学費捻出、家族の説得、…
- ◆ そこで研究・教育に興味があれば大学教員へ(もしくは西海岸へ)



前職でやっていたこと



- 学生（京大）の頃にやっていたこと
 - ◆ 京大情報学研究科の河原達也助教授（当時）の研究室で音声認識(HMM音響モデル)の研究 → 春の音響学会で発表

- 前職（京大研究員）でやっていたこと
 - ◆ 出戻りで同じ研究室で音声認識の研究開発
 - › 衆議院審議音声書き起こし、議事録作成支援
 - » <https://kokkai.ndl.go.jp/>
 - › 対話ロボットのフロントエンド
 - » 雑音下ストリーミング音声認識
 - › コーパス作成
 - » 雑音下話し言葉音声データ
 - › 低資源言語の音声言語処理
 - » アイヌ語音声認識合成

博士進学について（時期や進学理由）



- 時期 2021年4月
 - ◆ 同じ大学で働きながら社会人博士になるのは不可とのことで、普通の博士課程の学生として入学

- 理由
 - ◆ 2020年でちょうど研究員の雇用年限が来て、半年の暇ができた
 - ◆ 半年経てば同条件で再雇用可能だが、その間遊んでるくらいなら学生をやったほうがマシとのことで、入学
 - › 結局、RA等の賃金が発生しない半年間は完全に遊んで暮らすことに
 - 学位を取るのに1年半かかる、ただし無収入なので学費全額免除となり助かった
 - ◆ 学位があれば（プロジェクトつきのアルバイトではなく）正規の研究員として雇用可能らしいから

転職



- 前史
 - ◆ 国会審議音声議事録システムをNTTと京大で共同開発
 - › 当時は毎年政権が変わっておりシステムの定期更新が重要だった
 - ◆ 成果を2011年3月11日春季音響学会で発表→被災→(やむを得ず)朝まで語り明かす
 - ◆ 他にもNTTには知り合い（京大のときの学生さん含め）がたくさんいた

- 学位取得のタイミングで条件や入社後の仕事について伺い1年後転職
 - ◆ 仕事：雑音下のストリーミング音声認識
 - › 京大のときにやっていたことを継続
 - ◆ **+居住地の自由**
 - › 京都に住み続けられることが重要な条件だったため
 - › 時差のない場所だったらどこでもいいらしいが…
 - ◆ 京大との共同研究も開始

大学と企業のちがい



- 大学（京大）
 - ◆ 公共的な知(論文やソフトウェアや国のプロジェクト)を生産すること自体が目的
 - ◆ 一切の管理なし、勤怠も好きなときに紙にハンコを押すだけ、研究室にいつまでいてもいい
 - ◆ 旅費や学会参加費等も立て替えと事後の請求書精算が基本、物品購入なども口頭で依頼

- 企業（NTT）
 - ◆ 技術の「出口」が問われる（場合がある）
 - › 論文でなく、製品やサービスのこと
 - » NTTにも基礎研究所はある
 - ◆ 「マネジメント」の存在
 - › 勤怠、進捗管理、健康、安全、…
 - › 勤務時間外のサーバアクセス厳禁
 - » 手厚いサポートとも言える
 - ◆ 「事務」の時間が増えた（そうは言っても大した事ない）
 - › 起票→承認→決済
 - » この発表 → 関大前への旅費申請起票、このスライドの投稿許可願(一度否決される)

大学と企業のちがい 研究スタイル



- 大学（京大）
 - ◆ 教授を頂点とする垂直構造
 - ◆ 何をしても良い
 - › ただし、一応プロジェクトの仕事もしないといけない
 - ◆ いちばんの良さは、若い人と日々交流できる点
 - ◆ MacBookから四六時中サーバにアクセス

- 企業（NTT）
 - ◆ 数多の「同僚」から構成される水平構造
 - › いちばん気に入っているポイント
 - ◆ 何をしても良い（多分）
 - › そうはいつでも出口を意識するようになる
 - › 誤り率や遅延の削減にも興味が出てくる
 - ◆ チームの仕事が基本
 - › 自分のやることが他の人の仕事や生活に関わる(気がする)
 - › モデルを作ると、それを製品上で動く形式に変換して動作確認して性能評価してくれる人がいる…
 - ◆ Windows一択

日々の活動サイクル



■ 原則フルリモート

◆ 先週の活動

- › 月：同僚とオンラインでコーディング2時間、火：グループ会議2時間、開発打ち合わせ 1 時間、水：同僚と研究打ち合わせ 3 時間、開発会社と会議 1 時間、木：同僚と論文打ち合わせ 1 時間、金：有休、
- › それ以外は自由：論文(読む、書く)、コーディング、査読、オンライン雑談、…
- › その前は10日間夏休み

◆ cf. 京大時代：

- › 月2くらいのグループ（音声認識）ミーティング、年2くらいのプロジェクトミーティング
- › 基本的に、毎日夏休み

■ 月に二度ほど横須賀（+横浜、品川）に出社（1回2泊ほど）

- ◆ マシンや機材のセットアップ、防災訓練、データ収録、組織ぐるみの会合
- ◆ その他、グループメンバーとの密な「打ち合わせ」

音響学会に対するスタンス



■ 前職 京都大学

- ◆ 音声認識の学生には必ずおすすめした
 - › 研究会より聴衆が多くて修行になる + 奨励賞がある
 - › 修士は一応外部投稿ありが必須条件なので、早い時期にASJでその基準を満たしておくのが精神衛生上良い
 - › +日本は音響（音声）の研究のレベルが妙に高いので勉強になる

■ 現職

- ◆ ウェイト高い
 - › ASJに出すかどうかの確認が回ってくる
 - › 委員をしている社員もたくさんいる
- ◆ 春はInterspeech、秋はICASSPの締め切りと連動してるので良い
 - › 例えば若い人には1 2月の音声言語シンポジウムや3月のリゾート研究会とともにおすすめ
- ◆ 締め切りが早い

まとめ



■ 音響研究業界について（特に音声A）

- ◆ 見てもらったようなダラダラした生き方が許容されるほど懐が深い
- ◆ 音声は食いっぱぐれない
 - › 系列データで入力が連続値なので難しい
 - › 右の写真の収録データを認識するとまだ誤り率30%くらい
 - » ただしILRMAで15%ほどになる
 - › 行くところ行くと新卒二年目で年収8,000万だとか



■ 大学と企業のちがい

- ◆ 色々書いたが、基本的に「ない」
 - › 写真のようにマイクがあれば成立する
- ◆ おそらく、両方に所属するのがいちばん

前職の労働風景

現職の労働風景



■ 若い人々へのおすすめ

- ◆ 研究を仕事にするなら、結局のところ学位はとっておくのがよい