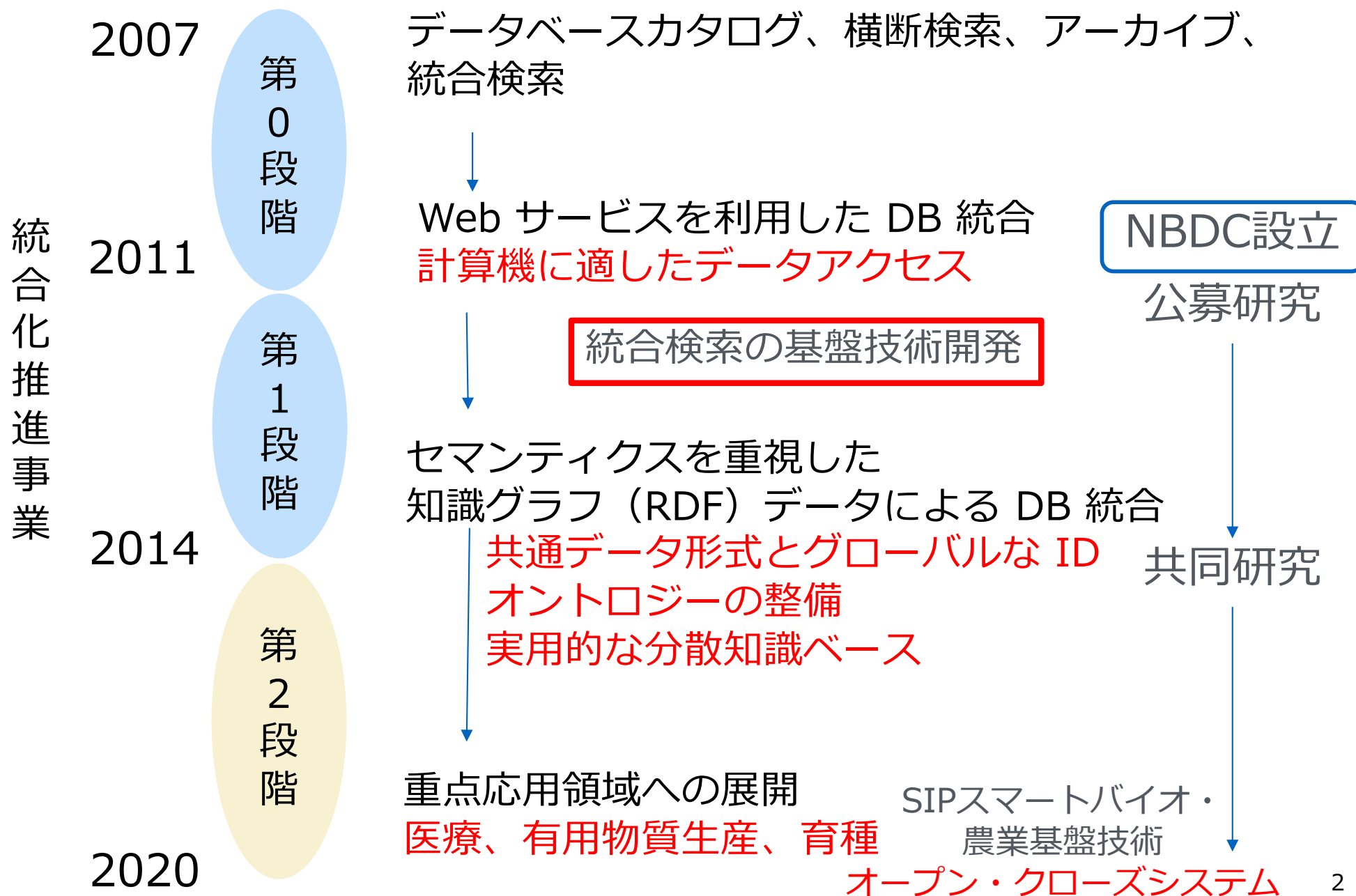


DBCLSにおける基盤技術開発

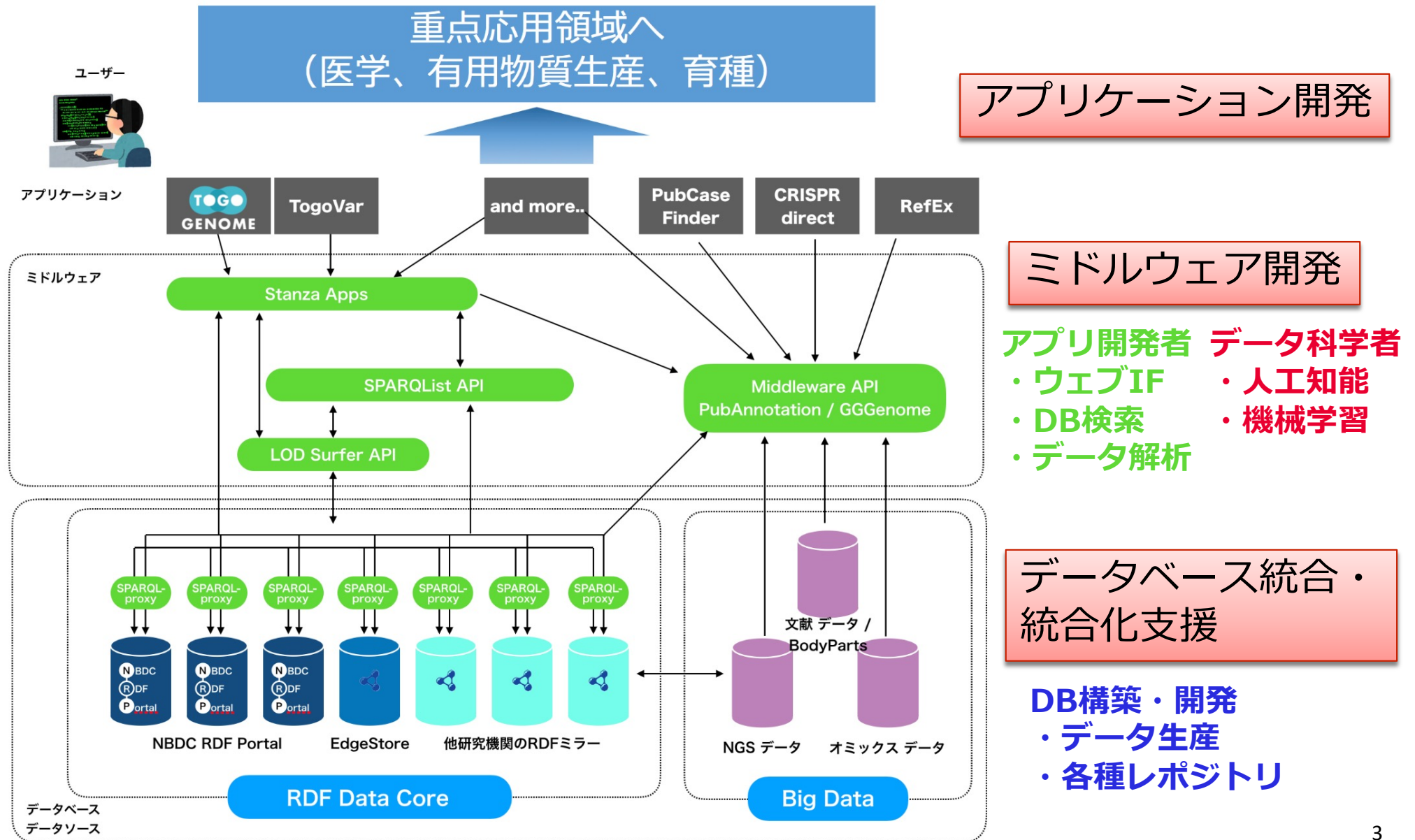
情報・システム研究機構
データサイエンス共同利用基盤施設
ライフサイエンス統合データベースセンター(DBCLS)

五斗進

DBCLSの活動



RDF データコアによる知識グラフを中心とした 研究開発



データベース統合・統合化支援

- データベースの標準化・RDF 化
 - RDF 化ガイドライン
 - RDF-config : 複雑なグラフデータの理解
 - TogoWS, TogoDB, D2RQMapper
 - TogoID : ID変換、関連情報の取得 (P3)
 - オントロジー構築 : 難病 (P19)、生物資源 (P40)
 - NBDC RDF ポータルへのデータ登録
 - SPARQLthonを通じた技術支援
 - 統合化推進プログラム
 - SIP : スマートバイオ・農業基盤技術
- 知識 (テキストデータ) の標準化・RDF 化
 - PubAnnotation : 文献のアノテーションデータを蓄積
 - COVID-19 関連の文献アノテーション (ハッカソン)

TogoID: DB統合の基盤となるID変換サービス



63データベース
156データベースペア

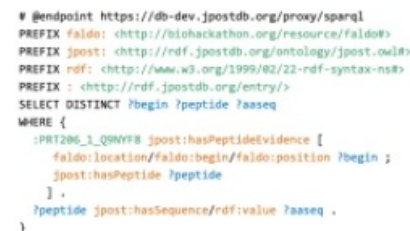
遺伝子・タンパク質・
化合物・パスウェイ・
疾患など

データベースアクセスミドルウェア

- データベースクエリ（SPARQL）作成・データアクセスの効率化
 - クエリの再利用：SPARQLList, SPANG
 - クエリ作成支援：SPARQL support, EndpointBrowser, LOD surfer
 - SPARQL エンドポイント確認：UmakaData, UmakaView
 - 可視化：TogoStanza、MetaStanza (P4)
 - データアクセス効率化：SPARQL-proxy
- DB開発プロジェクトで応用
 - TogoVar, TogoGenome
 - 統合化推進プログラム：MicrobeDB.jp, jPOST, GlyCosmos
 - SIP：スマートバイオ・農業基盤技術
 - 製品評価技術基盤機構：DBRP、微生物資源と関連情報

SPARQL support web ブラウザ上で動作する SPARQL クライアント・SPARQL クエリエディタ

視覚的なネットワークグラフを用いてSPARQLエンドポイント内のRDFデータを探索するwebサービス。
ノードをクリックしていくことで、SPARQL クエリを発行。



アプリケーション

➤ TogoVar: 日本人ゲノムの変異情報 (P22)

- 変異のアノテーション情報
- GWASによって得られた結果
- 疾患と変異との関連
- 論文と変異との関連

☒		WGS	GEM-J WGA	78,767,666
☒		WES	JGA NGS	4,561,277
☒		SNP	JGA SNP	1,249,723
☒		WGS	ToMMo 4.7KJPN	65,598,014
☒		WES	HGVD	554,461
☒		WES	ExAC	9,107,451
☒		Disease	ClinVar	663,441

➤ PubCaseFinder: 症状と希少疾患と遺伝子 (P25)



PubCaseFinder

文章から症状を自動抽出したい場合はコチラ

➤ TogoGenome: ゲノムと関連情報

➤ TogoMedium: 微生物培養培地情報

コミュニティ戦略

- ハッカソンによるデータベース開発者などの国際・国内連携
 - 国際版・国内版バイオハッカソン
 - Elixir BioHackathon
 - Biomedical Linked Annotation Hackathon
 - SPARQLthon : 毎月開催
- 学会展示、講習会
- TogoTV (P53)
 - 動画マニュアル
 - 講演・講習会
 - イラスト

DDBJ との連携

- ワークフローのDDBJスパコンでの実装 (P6)
 - GA4GH WES API 標準の DDBJ WES
 - DDBJ ヒトゲノム標準解析パイプライン
 - NGS 解析パイプライン
- 発現データ : RefEx, AOE
- 高速ゲノム検索 : GGGenome / CRISPRdirect (P13)
- データベース中の遺伝子アノテーション : TogoAnnotator (P9)

まとめと課題

- データベース統合化のための基盤技術
 - 統合化のための仕組みやデータセットは揃ってきた
 - 応用に向けた枠組みづくり
 - TogoDX
- バイオデータ科学者・機械学習応用との連携
 - TogoDXを用いたデータセットの取得
- 人材育成：草の根的に
 - ハッカソンなどによる人材交流
 - TogoTVなどによる教育素材

DBCLSのスタッフ

<http://dbcls.rois.ac.jp/members.html>



センター長
小原 雄治 / KOHARA, Yuji, PhD
 ゲノム科学
[Mail](#) [ORCID](#)



教授 (副センター長)
五斗 進 / GOTO, Susumu, PhD
 データベース統合
[Mail](#) [ORCID](#) [Google Scholar](#)



教授 (兼務)
大久保 公策 / OKUBO, Kousaku, MD, PhD
 ヒト統合、ポータル、発現統合、基盤技術
 担当サービス: BodyParts3D/Anatomography
[Mail](#)



特任准教授
片山 俊明 / KATAYAMA, Toshiaki, PhD
 セマンティック・ウェブ技術によるゲノム情報の統合化、
 担当サービス: TogoWS, TogoDB, TogoGenome, TogoSt
[Mail](#) [GitHub](#) [ORCID](#) [researchmap](#)



特任准教授
川島 秀一 / KAWASHIMA, Shuichi, PhD
 セマンティック・ウェブとバイオインフォマティクス技術
 担当サービス: NBDC RDF Portal, TogoVar, TogoGenom
[Mail](#) [GitHub](#) [ORCID](#) [researchmap](#)



特任准教授
金 進東 / KIM, Jin-Dong, PhD
 自然言語処理技術を利用したデータベースの生産性と接点
 担当サービス: PubAnnotation, PubDictionaries, LODQA
[Mail](#) [GitHub](#) [ORCID](#) [researchmap](#)



特任准教授
畠中 秀樹 / HATANAKA, Hideki, PhD
 生命科学系データベースの統合的利活用
[Mail](#) [ORCID](#)



特任准教授・DSコーディネーター
箕輪 真理 / MINOWA, Mari, PhD
 広報&もろもろの渉外窓口
[Mail](#)



特任准教授
山本 泰智 / YAMAMOTO, Yasunori, PhD
 テキスト処理とセマンティックウェブ技術を用いたデー
 担当サービス: inMeXes, Allie, Colil, LSD RDF Data Port
[Mail](#) [GitHub](#) [ORCID](#)



特任助教
大田 達郎 / OHTA, Tazro, PhD
 大規模データ解析に資するデータベース技術開発と塩基配列情報
 担当サービス: DBCLS SRA
[Mail](#) [GitHub](#) [ORCID](#)



特任助教
小野 浩雅 / ONO, Hiromasa, PhD
 「統合TV」編集・管理、遺伝子発現等の大規模データ解析および
 担当サービス: TogoID, TogoTV, RefEx
[Mail](#) [GitHub](#) [ORCID](#) [researchmap](#) [Google Scholar](#)



特任助教
千葉 啓和 / CHIBA, Hirokazu, PhD
 データベース統合化のための技術開発、およびデータベース利用
 担当サービス: SPANG
[Mail](#) [ORCID](#)



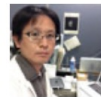
特任助教
内藤 雄樹 / NAITO, Yuki, PhD
 塩基配列などの大規模データを活用する技術の開発
 担当サービス: GGGenome, CRISPRdirect, GGRNA
[Mail](#) [GitHub](#) [ORCID](#) [researchmap](#)



特任助教
仲里 猛留 / NAKAZATO, Takeru, PhD
 NGSデータ利活用促進、オミックスデータの生物学的な解釈のた
 担当サービス: Gendoo, DBCLS SRA
[Mail](#) [GitHub](#) [ORCID](#) [researchmap](#) [Google Scholar](#)



特任助教
藤原 豊史 / FUJIWARA, Toyofumi, PhD
 希少・遺伝性疾患に関わるソフトウェア開発およびデータ整備
 担当サービス: PubCaseFinder, Allie, Colil, LSD RDF Data Port
[Mail](#) [GitHub](#) [researchmap](#) [Google Scholar](#)



特任助教
守屋 勇樹 / MORIYA, Yuki, PhD
 基盤技術
 担当サービス: Endpoint browser, SPARQL support
[Mail](#) [ORCID](#) [Google Scholar](#)



特任研究員
申 在紋 / SHIN, Jae-Moon, PhD
[Mail](#)



特任研究員
三橋 信孝 / MITSUHASHI, Nobutaka
 バリエーションを中心としたヒトデータの統合、サーバ管理
 担当サービス: TogoVar
[Mail](#) [GitHub](#) [ORCID](#)



学術支援技術専門員
王 悦 / WANG, Yue, PhD
 テキストマイニング
 担当サービス: PubAnnotation
[Mail](#)



学術支援技術専門員
池田 秀也 / IKEDA, Shuya
[Mail](#)



学術支援技術専門員
高月 照江 / TAKATSUKI, Terue
[Mail](#)



学術支援技術専門員
平木 愛子 / HIRAKI T., Aiko
 アノデータ
[Mail](#)



学術支援専門員
佐久間 桂子 / SAKUMA, Keiko
 広報普及、企画立案、運営管理など
[Mail](#)



准教授(兼務)
川本 祥子 / KAWAMOTO, Shoko, PhD
 データベースと文献の検索、辞書の構築と公開、オープンアクセスなど
[Mail](#)



特任准教授
岡本 忍 / OKAMOTO, Shinobu, PhD
 セマンティック・ウェブ技術による生物メタデータの標準化とアプリケーションのデザイン
[Mail](#)



開発支援(サーバ管理)
川越 千晴 / KAWAGOE, Chiharu
 開発支援 (サーバ管理)
[Mail](#)



開発支援(サーバ管理)
小笠原 雅喜 / OGASAWARA, Masaki
 開発支援 (サーバ管理)
[Mail](#)