



異分野融合・新分野創成を担う データサイエンティストの 育成基盤

統計思考力を備えた人材育成による融合研究の推進を目指して

情報・システム研究機構シンポジウム

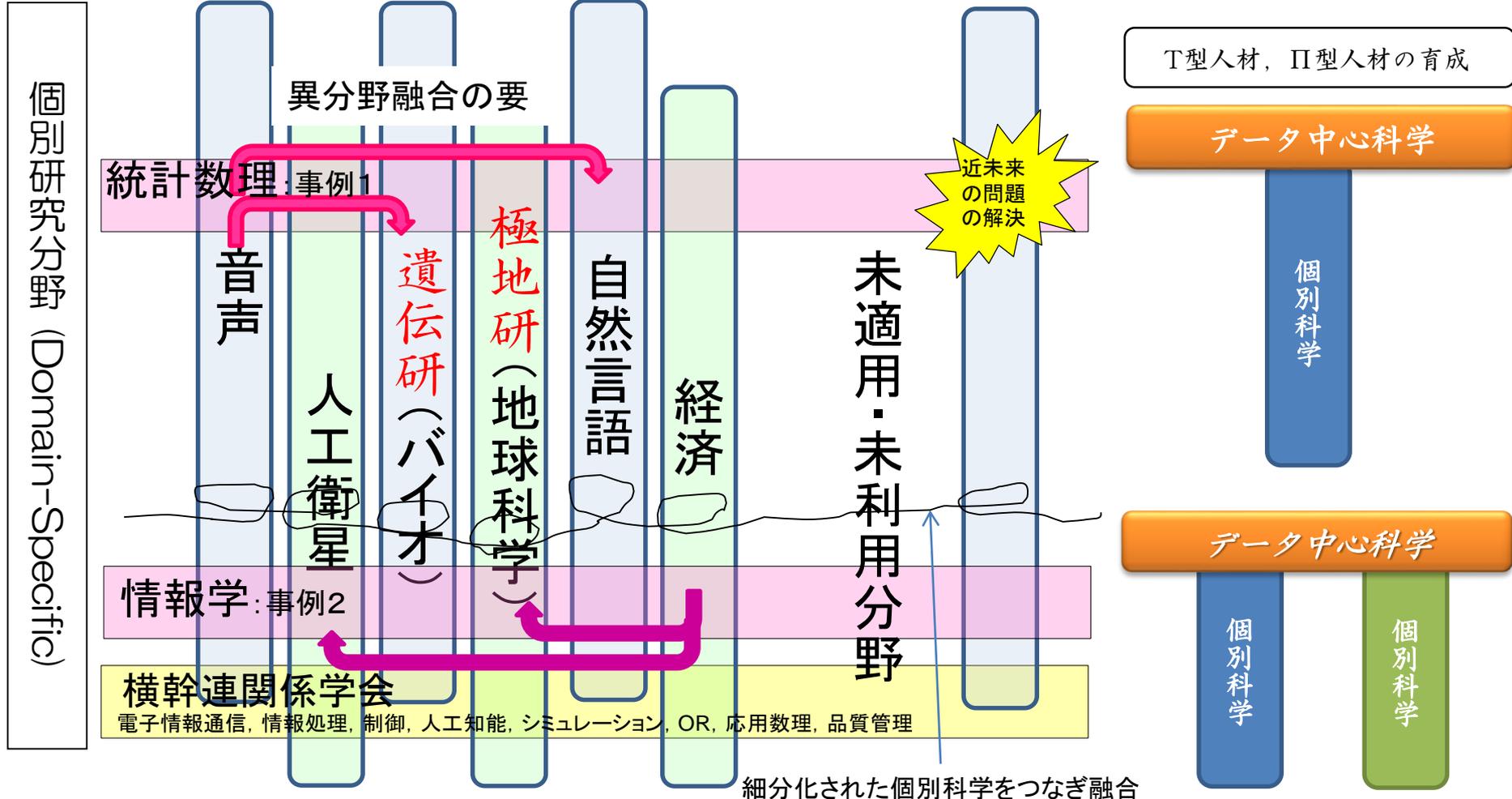
2017年2月20日(月)

情報・システム研究機構
統計数理研究所 樋口 知之



T型・II型人材の育成

横断型研究分野 (Domain-General)



情報・システム研究機構全体の取組み

法人第一期 (2004-2009)

機構内中心の融合研究と生命科学分野のDB統合化を推進

ライフサイエンス
統合DBセンター

新領域融合
研究センター

法人第二期 (2010-2015)

データサイエンスの三位一体の基盤整備事業を推進

データ中心科学
リサーチコモンズ事業

ライフサイエンス
統合DBセンター

新領域融合
研究センター



法人第三期 (2016-2021)

支援事業（データ共有、データ解析、人材育成等）を中心とした活動を積極的に推進

データサイエンス
共同利用基盤施設

データサイエンス支援

- データ共有支援
- データ解析支援
- T型・II型人材育成

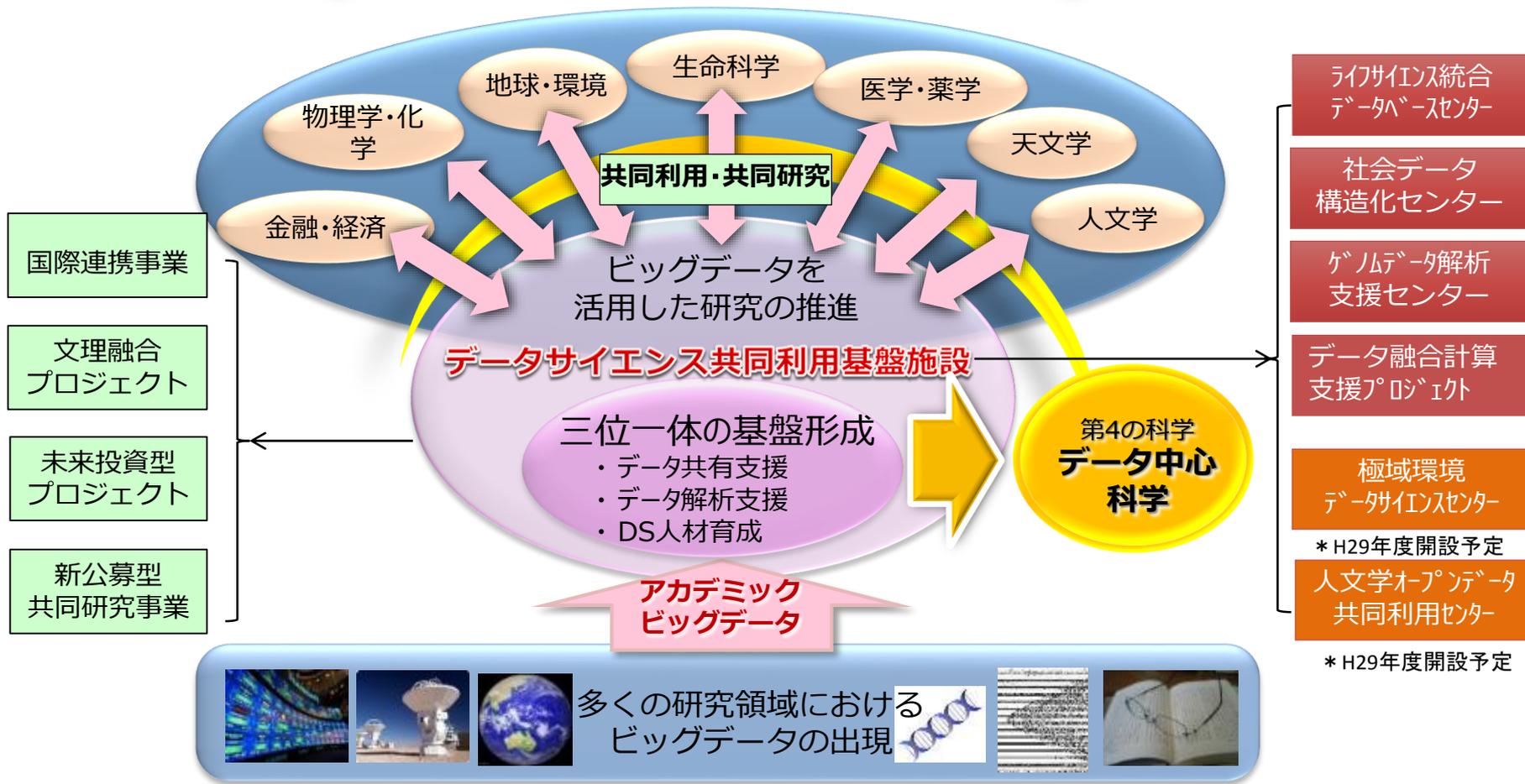
戦略プログラム

- 戦略的研究ネットワーク形成
- 未来投資型研究プログラム

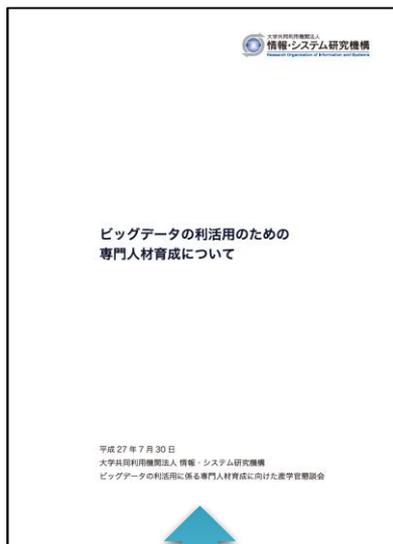
データサイエンスの観点から大学等の研究活動を支援

データサイエンス共同利用基盤施設

分野を超えた知識移転・汎化、異分野交流・新分野創成を実現



【提言】ビッグデータの利活用のための専門人材育成について

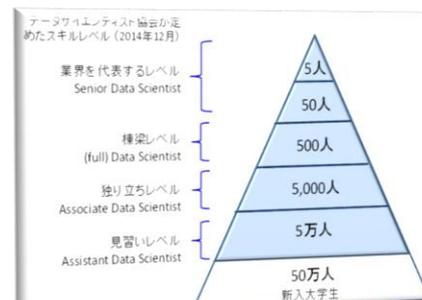
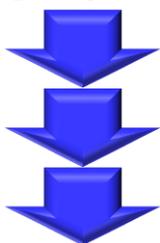


我が国の問題の根源は、**棟梁レベルの決定的不足**にある。この解決のために**国家レベルの拠点を設置**して、年500名規模の「棟梁レベル」の人材育成をめざし、上層への成長や下層へのトリクルダウン効果も狙う。

リテラシーレベルや独り立ちレベルの**大学教育を加速させる**ために、主要10大学程度で本報告書の提案に基づく人材育成をスタートすると共に、MOOCなどのオンライン教材を整備し、全国への波及効果を狙う。

社会全体のリテラシーやアウェアネスを向上させるために、全学的教養教育の実施、国家レベルのフラッグシップ・プロジェクトの推進、コンテストの開催、映像素材の充実などの取組を行う。

大学共同利用機関法人
情報・システム研究機構



これらの方策を実現するにあたっては、データサイエンスを副専攻とする**ダブルディグリー制****人材育成の推進**や**スキル認定制度も有効**と考えられる。

若手研究者育成：

「ROIS/I-URIC 若手研究者クロストーク」

①目的等：異分野融合と交流の機会創出。
合宿形式。

今年度は講演・研究紹介・討議
の全てを英語で実施。

②対象等：4機構の若手研究者＋学生。
総研大との共同開催。

(H28実績：参加総数50名：
ROIS 32名、3機構8名、
総研大生10名)

③平成28年度テーマ：「分かり合えるコミュニケーション」



国際開発者会議BioHackathon

ライフサイエンス統合データベースセンター（DBCLS）では、生命科学分野の統合データベース構築のための技術基盤の確立を目的とし、国際開発者会議BioHackathonを開催している。

最先端のデータ統合技術やその応用についての世界の技術動向を調査・把握しテーマの決定を行い、国内外の先端技術を用いてシステムやプログラム開発を行っている現場の研究者を参集して開催。



合宿形式で分野横断的に問題解決を行い、密度の高い情報交換と生産性の高い技術開発が会議の場でリアルタイムに行われる点
が大きな特徴となっている。



DBCLS はこの分野で世界をリードするハブとして高く評価され、今後も継続的にリーダーシップを発揮することが期待されており、その成果はDBCLSの開発する統合データベースの基盤技術として活用されているほか、研究者コミュニティの共有資産となり論文発表も行われている。

統合データベース講習会：AJACS

(All Japan Annotator/Curator/System DB administrator)

統合データベース講習会：AJACSは、生命科学系のデータベースやツールの使い方、データベースを統合する活動を紹介する初心者向けのハンズオン講習会。

2007年度からDBCLS
が主催。

2011年度からは
JST/NBDCが主催し、
DBCLSは共催。

これまで60余回、全国
各地で開催。



SPARQLthonの開催

- ・ H24年10月より、**セマンティックウェブ技術**による**生命科学DBの統合**をテーマに、毎月DBCLS主催で開催しているハッカソン。
- ・ これまで参加延べ人数は**975人**、ユニークな参加者数は**112名**、所属機関数は**39**（12大学、13研究機関、14企業）。
- ・ H26年度からは、統合化推進プログラムから多くの方に参加頂いており、データのRDF化に関する**意識共有・技術情報共有・コラボレーション**が著しく進んだ。また、ここで多くのRDFデータやオントロジーが作られてきた。



H29年1月時点で、SPARQLthon 51まで開催。
第一回からの累積では、参加者1,275人。

統計思考院

● 統計数理研究所の今期中期目標・中期計画にある「**統計思考力を備えたT型人材育成による融合研究の推進**」を実現するため、新しい統計学の創成を目指す研究者や学生、固有分野の研究で統計学の必要性を感じた人、などさまざまな人が集い切磋琢磨しながら「統計思考」の訓練や研究をする

「統計思考 “道場”」



期待される効果・目的

1. 大規模データを活用したモデリングや研究コーディネーションなど大規模データ時代に求められる**統計思考**ができる人材（T型人材，モデラー，研究コーディネータ）を共同研究の現場で育成
2. 大学教員のサバティカル支援事業や統計数理の成果の**公開普及活動**の実施
3. 複雑・不確実な現象に挑戦する**新しい研究分野の創成**

◆ 右：統数研の教員による統計数理の最新トピックを取り上げる「統計数理セミナー」
⇒研究者修業の一環として、発表者からの事前ヒアリングを経て、**統計思考院特任助教**が打合せ、司会を交代で行う

◆ 左：統計思考院セミナー
⇒**統計思考院特任助教**が企画開催

共同研究スタートアップ

- 問題の適切な位置付けや，課題解決のための相談に対応
- 研究者，その他の方々に統計数理に関わる問題，データ解析・統計分析等でお悩みの方に統計専門家がその難易を見極め，解決に向けて助言を行うプログラム
 - ✓ **相談の内容によって，共同研究へ発展する場合も有**



On the Job Training (OJT) として特任助教2名が特命教授 (メンター) 2名につき，ともに対応することで計思考力，統計数理科学およびその応用・方法論等の知識を深化



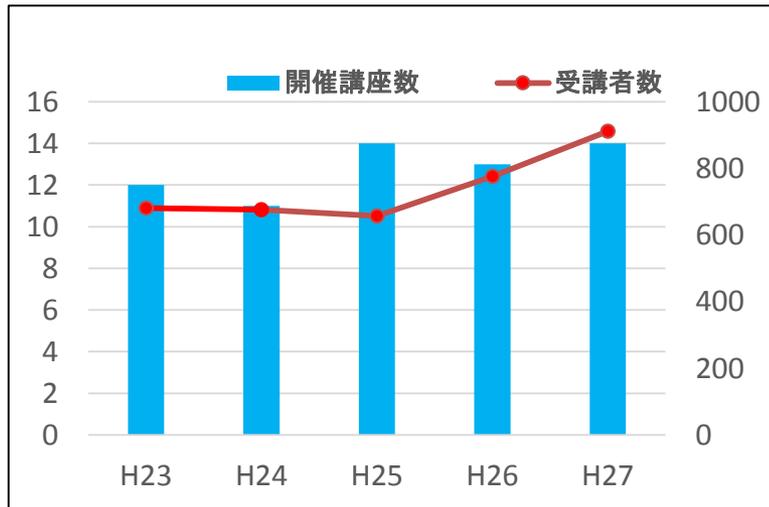
共同研究スタートアップ申込用紙	
2024年 月 日 ()	氏名
お問い合わせ先	〒100-8302 東京都千代田区千代田 1-9-1 統計数理研究所 共同研究推進課
連絡先	TEL 03-5462-3111 FAX 03-5462-3112
お問い合わせ内容	共同研究スタートアップの申し込みについてお尋ねください。
お問い合わせの件名	
お問い合わせの経緯	
お問い合わせの目的	
お問い合わせの連絡先	
お問い合わせの電話番号	
お問い合わせのメールアドレス	
お問い合わせの住所	
お問い合わせの郵便番号	
お問い合わせの市区町村	
お問い合わせの支庁	
お問い合わせの都道府県	
お問い合わせの国	
お問い合わせの性別	
お問い合わせの年齢	
お問い合わせの学歴	
お問い合わせの職歴	
お問い合わせの専攻	
お問い合わせの学位	
お問い合わせの学位取得年	
お問い合わせの学位取得機関	
お問い合わせの学位取得科目	
お問い合わせの学位取得単位	
お問い合わせの学位取得成績	
お問い合わせの学位取得状況	
お問い合わせの学位取得予定	
お問い合わせの学位取得予定年	
お問い合わせの学位取得予定機関	
お問い合わせの学位取得予定科目	
お問い合わせの学位取得予定単位	
お問い合わせの学位取得予定成績	
お問い合わせの学位取得予定状況	
お問い合わせの学位取得予定予定	
お問い合わせの学位取得予定年	
お問い合わせの学位取得予定機関	
お問い合わせの学位取得予定科目	
お問い合わせの学位取得予定単位	
お問い合わせの学位取得予定成績	
お問い合わせの学位取得予定状況	
お問い合わせの学位取得予定予定	



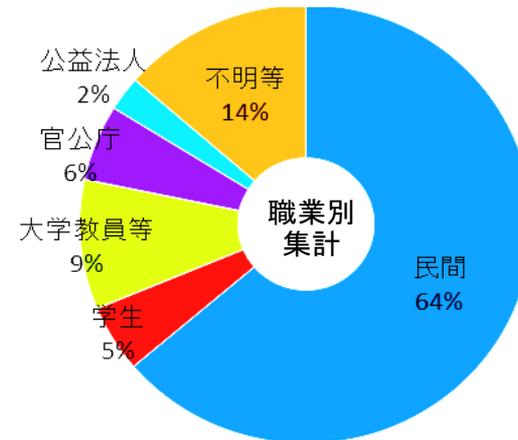
公開講座



H27年度14講座開催
総受講者数：911人



H28年度
前期：7講座開催後期：7講座開催



平成27年度受講者内訳
* 8割以上が社会人受講者

組織連携 (MOUや組織長同士の合意) に基づくデータサイエンス講座企画

一例

人工知能／ビッグデータ／IoT／サイバーセキュリティ統合プロジェクト
【中核機関：理化学研究所・革新知能統合研究センター (AIP)】に関する人材育成事業

平成28年度「機械学習速習コース」

対象：若手研究者・大学院生・企業研究者等
(講義内容は、理化学研究所 革新知能統合研究センターとの連携・共同研究等を行う水準に設定)

* 関係者のみ参加可能。申し込み受付開始から3日で満席



2月9日 (木) 受付停止	A M	池田 思朗	「スパース性を用いた情報処理」
	P M	江口 真透	「パターン認識のための機械学習」
3月7日 (火) 受付停止	A M	吉田 亮	「マテリアルズインフォマティクス：機械学習による新物質探索」
	P M	持橋 大地	「離散データの統計的機械学習」
3月28日 (火) 受付停止	A M	武田 朗子	「不確実性を考慮した最適化法と機械学習への応用」
	P M	福水 健次	「カーネル法による非線形データ解析」

e-learning教材作成

Scorm 1.2形式で作成、効果測定用にトピックごとにタグ付け

- ▶ 統計数理研究所運営会議の外部委員や統計関連学会関係者からのご提案を受けて、平成18年（2007）度より全国の大学院生および学生のための夏の学校を開催しており、統計数理に係るテーマを一つ取り上げ、1～2日間の講義を開催してきたもの



平成26（2014）年度からは、公募型人材育成事業ワークショップとのコラボレーション企画 **連続10日間** のコースとして開催（いわば、**統計数理ブートキャンプ**）

情報・システム研究機構 統計数理研究所
統計思考力育成事業 夏期大学院

感染症流行の数理モデル

夏期短期（入門）コース開講 受講生募集

- 連続10日間（休日なし）のブートキャンプ
- 本研究分野におけるオールジャパニー一流の講師陣
- 外国からの著名な講師も招へい
- 実践プログラミングまでのスキル向上を目標とするコース

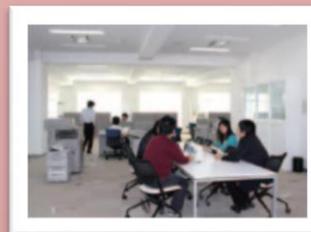
【参加実績】

H26年度 受講者80名 講師・フュータ-26名
H27年度 受講生86名 講師・フュータ-28名
H28年度 受講生78名 講師・フュータ-27名
*講師・フュータ-には外国人も含む



- 「受託研究員制度」を利用した産学官（特に企業）への共同利用提供
- 企業側のニーズに応じた人材育成や人材交流面でのフォロー
- 共同研究の基盤作り
- 平成27年5月より受入れを開始
- 平成27年度実績（3名受入れ）
 - キヤノン株式会社
 - 株式会社ブリヂストン
 - SAS Institute Japan 株式会社
- 平成28年度実績（4名受入れ）
 - 株式会社ブリヂストン
 - アズビル株式会社
 - 塩野義製薬株式会社
 - 株式会社 クニエ

◎受託研究員が利用可能な統数研内資源等



⇒毎週の「統計数理セミナー」をはじめとする多くの研究集会等を通じた統数研約50名の常勤教員，年間100名に近い国内外からの訪問研究者との交流



⇒初級～上級まで幅広いレベルの公開講座の受講



⇒統計学・数学・情報科学関係蔵書6万冊を誇るライブラリ，電子書籍，学術書誌情報の利活用



⇒世界最大級の共有メモリ型スーパーコンピュータシステムの利活用 * [HPCIの利用登録要](#)

トップレベルのIT技術者育成を目的とした教育プログラム

最新のソフトウェア
科学・工学技術



実際の開発現場で
活用するノウハウ



スーパーアーキテクト
(開発現場に最新の開発技術を導入できる人材)
を育成

講義受講

- ・産学連携で教材開発
- ・実践に近いリアルな例題で演習
- ・7シリーズ, 43科目



ソフトウェア科学・工学の基礎から
応用まで学べる科目群を提供

修了制作

受講生それぞれの開発現場における
実問題を分析・モデル化し、
最新のソフトウェア科学・
工学技術の導入によって解決

- ・3～6ヶ月実施
- ・講師がマンツーマンで指導し
ノウハウを伝授
- ・考案した解決を**実際の開発現場に
導入**し、効果測定
- 学んだ技術を実践で活かすための
ノウハウを修得する

実績

- ・毎年約40名の企業IT技術者を教育
- ・過去11年で348名の育成実績
- ・文部科学省 科学技術振興調整費
の支援(平成16～20年度)で立ち上げ
- ・平成21年度より自主事業として継続
- ・協賛企業54社との協力関係
(日立製作所, 富士通, 日本電気, 東芝
NTTデータ, キヤノン, 日本ユニシス等)
- ・講師40名(産業界25名, 学界15名)

<http://www.topse.jp>

トップエスイーコース

- ・基本となる1年間の教育コース (年額 557,280円)
- ・主にソフトウェア科学・工学の基礎技術修得に特化

アドバンス・トップエスイーコース

- ・1年間の上級教育コース (年額 1,080,000円)
- ・難度の高い先端的な開発課題の分析と解決に特化
(クラウド, IoT, ビックデータ, AI, ...)
- ・教科書化もされていない最先端技術・事例を共同で調査するゼミを設置

トップエスイー 実践プログラミングセミナーシリーズ

- ・先端プログラミング技術に特化した短期集中セミナー (1日8万円, 2日15万円)
- ・ビックデータ, AI, Blockchainsなどの先端トピックを時流に応じて実施

トップエスイーコンテスト

- ・先端的なソフトウェア開発技術を競うオンラインコンテスト

学術ネットワークSOCを通じた 技術職員の育成

■ 大学間連携に基づく情報セキュリティ体制の基盤構築

◆ 国立大学法人等の運営費交付金から拠出

◆ 3種類の監視システム

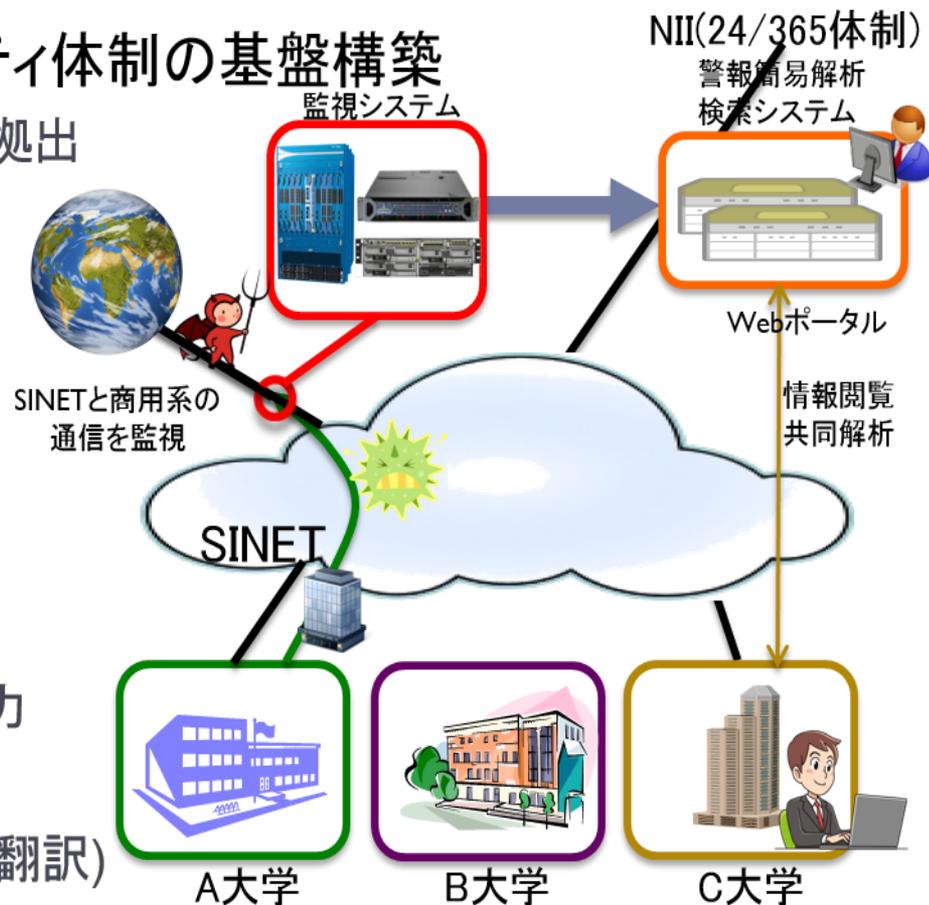
- Sandbox搭載IDS (paloalto)
- シグネチャベースIDS (Cisco FirePower)
- DNSトラフィック監視 (Damballa CSP)

◆ 簡易解析システム+Webポータル

- 膨大な警報に危険度・優先度付け

■ 橋渡し人材に求める人物像

- ◆ ネットワークの健康状態を俯瞰する能力
- ◆ 的確な状況把握と柔軟な判断力
- ◆ 経営目線(サイバー用語を経営用語へ翻訳)



学術研究機関へのベンチマークデータの提供 (2017年を予定)

■ 警報情報とセッション情報

◆ 一般公開

- IPアドレス無しの単なる統計データ
- Big data解析などでの活用

◆ NDAに基づく研究機関向け公開

- 大学の特定を回避する匿名化データ
- 未知の攻撃検知・攻撃対策の研究開発に活用

■ マルウェア情報

◆ NDAに基づく研究機関向け公開

◆ マルウェア本体+動作検証の解析結果

- 最新データを利用した新たな対策技術の開発を支援

学術研究・学生教育を支援



想定する活用例

白浜危機管理コンテストにおける人材育成
(橋渡し能力+技術力に関する学校対抗競技)

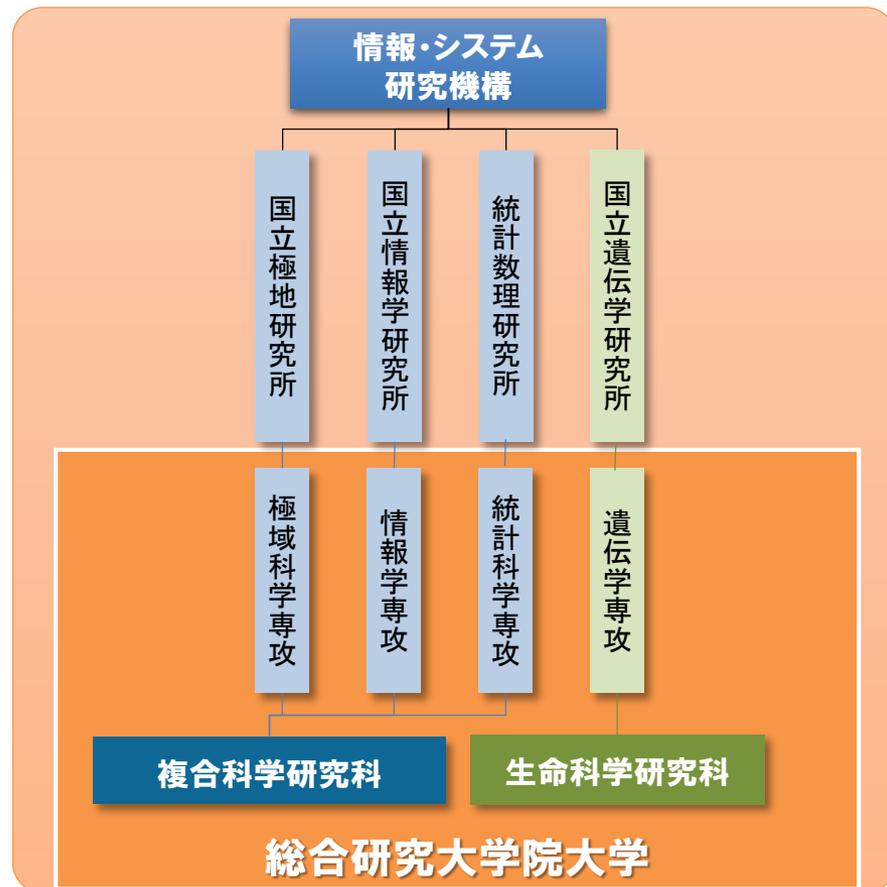
大学院教育に関する取組

総合研究大学院大学の基盤機関として

- 大学共同利用機関が有する優れた人材と研究環境を活用して博士課程の教育を行い、一流の研究者を養成する**総研大の基盤機関として大学院教育を実施。**
- 各機構長が**総研大の経営協議会**や**学長・機構長連絡協議会**に、また、大学共同利用機関等との連携を強化することを目的とした**総研大のアドバイザーボード**に各機構長が推薦する理事が参画するなど、**大学院教育の運営に積極的に関与。**

全国の国公立大学との連携

- 各大学共同利用機関は、全国の国公立大学から、**特別共同利用研究員等**として学生を受け入れ。
- 連携大学院等**を通じて各大学院独自の教育方針と大学共同利用機関の優れた研究環境が融合した多様な取組が行われており、その役割は重要。



平成27年5月1日現在	情報・システム研究機構
総合研究大学院大学	170
特別共同利用研究員等	31
連携大学院等	55